

Pengukuran Kemiripan Model Proses Bisnis untuk menentukan *Common Fragment*

Gandhis Ulta Abriani, Muhammad Ainul Yaqin, Fatchurrochman

Abstract— Common fragment process extraction is carried out to obtain general business processes, where business processes are always found in the business processes of every organization with the same scope. Calculation of the similarity value of business processes can be obtained based on structural similarity, and behavioral similarity using the jaccard coefficient similarity algorithm, and the formation of paths used in the behavioral similarity process using causal footprints, and semantic similarity using wu palmer, and Analytical Hierarchy Process. These three aspects of similarity are then sought for by their accumulation values by weighting. The similarity weight value is calculated using the Analytical Hierarchy Process method. The results of the accumulation of similarity values of business processes are used in the clustering process. The results of clustering are used in determining common fragment. The results showed that the calculation of the similarity of business process models occurred 45 times. The composition of structural, behavioral, and semantic weight values is (0.33, 0.10, 0.57). In the clustering process of similarity weighting values formed 3 clusters with a threshold value of 0.49. The results of the extraction of common fragments in the form of elements from semantic, structural, and behavioral are combined as a general process model that is modeled in the form of Business Process Modeling and Notation.

Keyword: *Business Process Modeling and Notation; Clustering; Ekstraksi Common Fragment; Similarity*

Abstrak—Ekstraksi common fragment process dilakukan untuk memperoleh proses bisnis umum, dimana proses bisnis tersebut selalu ditemui pada proses bisnis setiap organisasi dengan ruang lingkup yang sama. Perhitungan nilai kemiripan proses bisnis dapat diperoleh berdasarkan structural similarity, dan behavioral similarity menggunakan algoritma jaccard coefficient similarity, dan pembentukan alur yang digunakan pada proses behavioral similarity menggunakan causal footprint, dan semantic similarity menggunakan wu palmer, dan Analytical Hierarchy Process. Ketiga aspek kemiripan tersebut, kemudian dicari nilai akumulasinya dengan melakukan pembobotan. Nilai bobot kemiripan dihitung

menggunakan metode Analytical Hierarchy Process. Hasil akumulasi nilai kemiripan proses bisnis digunakan pada proses clustering. Hasil dari clustering digunakan dalam penentuan common fragment. Hasil penelitian menunjukkan bahwa perhitungan kemiripan model proses bisnis terjadi sebanyak 45 kali perhitungan. Komposisi nilai bobot structural, behavioral, dan semantic sebesar (0.33, 0.10, 0.57). Pada proses clustering dari nilai pembobotan kemiripan terbentuk 3 cluster dengan nilai threshold 0,49. Hasil ekstraksi common fragment berupa elemen-elemen dari semantic, structural, dan behavioral dijadikan satu kesatuan sebagai model proses umum yang dimodelkan dalam bentuk Business Process Modeling and Notation.

Kata kunci: *Business Process Modeling and Notation; Clustering; Ekstraksi Common Fragment; Similarity*

I. PENDAHULUAN

Proses bisnis merupakan unsur penyusun dari sebuah ERP (*Enterprise Resource Planning*). ERP merupakan sistem informasi yang diperuntukkan bagi perusahaan manufaktur maupun jasa yang berperan mengintegrasikan dan mengotomasikan proses bisnis yang berhubungan dengan aspek operasi, produksi maupun distribusi. Masing-masing perusahaan pada umumnya memiliki suatu ERP yang sesuai dengan proses yang ada pada perusahaan manufaktur tersebut. Hal ini akan membentuk sistem informasi tunggal yang hanya dikelola dan digunakan oleh masing-masing perusahaan. Kelemahan dari sistem informasi tunggal adalah ketika terdapat perubahan proses bisnis, maka pihak pengembang akan memerlukan waktu untuk menyesuaikan proses bisnis yang sudah ada dengan proses bisnis yang baru [1]. Meskipun pada setiap perusahaan memiliki proses bisnis yang berbeda, namun dalam ruang lingkup yang sama secara umum memiliki kemiripan beberapa aktivitas dalam proses bisnis tersebut. Dari kemiripan tersebut memungkinkan untuk memperoleh satu model umum. Proses bisnis yang digunakan sebagai inputan direpresentasikan menggunakan pemodelan BPMN (*Business Process Modelling and Notation*). Perhitungan kemiripan antar model proses berdasarkan kemiripan struktural (*Structural Similarity*), kemiripan perilaku (*Behavioral Similarity*), dan kemiripan semantic (*Semantic Similarity*) [1]. Setelah dilakukan perhitungan kemiripan, dilakukan proses *clustering*. Tahap yang selanjutnya adalah pembuatan *common fragment* yang merupakan hasil dari elemen yang terkluster.

Gandhis Ulta Abriani is with the Informatic Engineering Departement of Maulana Malik Ibrahim Islamic State University , Malang, Indonesia (email gandhisultaabriani10@gmail.com)

Muhammad Ainul Yaqin, was with the Informatic Engineering Departement of Maulana Malik Ibrahim Islamic State University, Malang, Indonesia (e-mail: yaqinov@ti.uin-malang.ac.id).

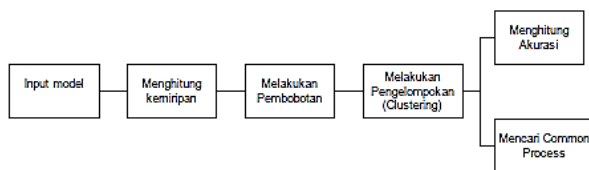
Fatchurrochman was with the Informatic Engineering Departement of Maulana Malik Ibrahim Islamic State University, Malang, Indonesia (fatchur@ti.uin-malang.ac.id)

Berdasarkan latar belakang tersebut dapat dilakukan penelitian untuk mendapatkan model proses bisnis yang umum. Pada penelitian yang telah dilakukan oleh Abdul Charis Fauzan adalah menentukan *behavioral similarity* menggunakan algoritma TARs (*Transition Adjacency Relations*) [2], sedangkan pada penelitian yang telah dilakukan oleh Arif Wahyu Prasetya adalah menentukan kemiripan *structural, behavioral, dan semantic*. Pada perhitungan kemiripan *semantic* mengacu berdasarkan label atau *string similarity*. Dalam menentukan bobot kemiripan dan nilai *threshold* dilakukan secara manual atau sistem uji coba [3]. Pada penelitian ini untuk menentukan kemiripan struktur dan perilaku menggunakan algoritma *Jaccard Coefficient Similarity*. Pada kemiripan *semantic* untuk menghitung nilai kemiripan makna katanya menggunakan WS4J dengan algoritma *Wu palmer* [4]. dan AHP (*Analytical Hierarchy Process*) untuk menentukan *sentence similarity*-nya. Setelah mendapatkan hasil perhitungan dari proses kemiripan, diharapkan dapat dilakukan penentuan *common fragment* model proses bisnis. Sehingga saat pencarian *common fragment* dapat dihasilkan model gabungan dari kemiripan proses bisnis yang telah ditentukan sebelumnya.

II. LANDASAN TEORI

Berdasarkan latar belakang tersebut dapat dilakukan penelitian untuk mendapatkan model proses bisnis yang umum. Pada penelitian yang telah dilakukan oleh Abdul Charis Fauzan adalah menentukan *behavioral similarity* menggunakan algoritma TARs (*Transition Adjacency Relations*) [2], sedangkan pada penelitian yang telah dilakukan oleh Arif Wahyu Prasetya adalah menentukan kemiripan *structural, behavioral, dan semantic*. Pada perhitungan kemiripan *semantic* mengacu berdasarkan label atau *string similarity*. Dalam menentukan bobot kemiripan dan nilai *threshold* dilakukan secara manual atau sistem uji coba [5]. Hasil dari penelitian tersebut bahwa algoritma *jaccard coefficient similarity* merupakan algoritma yang akurat untuk menghitung tingkat kemiripan model proses bisnis.

III. METODE PENELITIAN



Gambar 1. Prosedur penelitian

A. Input Model

Data yang digunakan sebagai inputan dimodelkan menggunakan BPMN kemudian di ekspor dalam file berekstensi *.xpd. Elemen-elemen yang terdapat pada BPMN mewakili alur proses yang berlangsung.

B. Menghitung Kemiripan

1) *Structural Similarity*

Structural similarity merupakan nilai kesamaan antar model proses bisnis dengan melihat dari struktur atau bentuknya. Menurut Remco Djikman [6] untuk menghitung kesamaan dua proses, harus ditemukan

pemetaan yang menginduksi kesamaan maksimal. Untuk menghitung nilai *structural similarity* dengan menggunakan persamaan *jaccard coefficient similarity*:

$$Sim (A,B)=\frac{A \cap B}{A \cup B} \tag{1}$$

A : Model proses 1 $A \cap B$: Jumlah yang sama
 B : Model Proses 2 $A \cup B$: Jumlah yang sama+Jumlah yang berbeda

2) *Behavioral Similarity*

Nilai *behavioral similarity* didapatkan dengan menghitung jarak kedua proses bisnis dalam ruang vektor yang dibangun dari jejak kaki kausal (*causal footprint*) kedua proses bisnis tersebut [7]. Dari jejak-jejak *causal footprint* yang terbentuk akan membentuk *node-node* yang akan saling terhubung. *Node-node* tersebut terbentuk dari *task* yang terdapat pada model proses yang sesuai dengan alur yang berjalan. Dari hasil penentuan *causal footprint* tersebut dapat dihitung nilai kemiripannya menggunakan persamaan *jaccard coefficient similarity* seperti pada perhitungan *structural similarity*.

3) *Semantic Similarity*

Perhitungan kemiripan semantik merupakan proses yang memerlukan keterlibatan beberapa aspek yakni tentang bahasa, komputer, matematika logik dan domain yang bersangkutan [8]. Perhitungan kemiripan *semantic* dihitung berdasarkan kemiripan label atau *string similarity* dan kemiripan *contextual* yang mengacu pada kemiripan makna antar kalimat. Pada perhitungan *contextual similarity* menggunakan algoritma *wu palmer* untuk menghitung kemiripan antar kata pada kalimat [8]. Kemudian menentukan nilai kemiripan kalimat dengan mengakumulasikan matriks nilai kemiripan kata dengan bobot nilai kelas kata yang dihitung menggunakan AHP. Kriteria yang digunakan yaitu *noun, dan verb*. *Noun*, dan *verb* merupakan kelas kata yang dihitung nilai kemiripannya pada WS4J. Di bawah ini merupakan persamaan untuk menghitung *sentence similarity* berdasarkan *contextual*-nya :

$$Sentence\ similarity = (\overline{n_{max}} \times b_n) + (\overline{v_{max}} \times b_v) \tag{2}$$

$\overline{n_{max}}$ merupakan nilai rata-rata dari kriteria *Noun*. $\overline{v_{max}}$ merupakan nilai rata-rata dari kriteria *Verb*. Sedangkan untuk b_n merupakan bobot nilai dari kriteria *Noun*. b_v merupakan bobot nilai dari kriteria *Verb*. Dari hasil perhitungan menggunakan AHP diperoleh komposisi nilai bobot kriteria *noun*, dan *verb* sebesar (0.25, 0.75).

C. Pembobotan Kemiripan

Pembobotan penghitungan kemiripan merupakan bobot nilai masing masing kriteria kemiripan yaitu *structural, behavioral, dan semantic* yang kemudian diakumulasikan menjadi satu dengan cara menjumlahkan hasil kali nilai kemiripan dengan masing-masing bobot kriteria kemiripan yang telah ditentukan. Bobot nilai kemiripan dihitung menggunakan algoritma AHP (*Analytical Hierarchy Process*).

Tabel 1. Nilai bobot kemiripan

Kriteria	Nilai bobot prioritas
<i>Structural</i>	0.33
<i>Behavioral</i>	0.10
<i>Contextual</i>	0.57

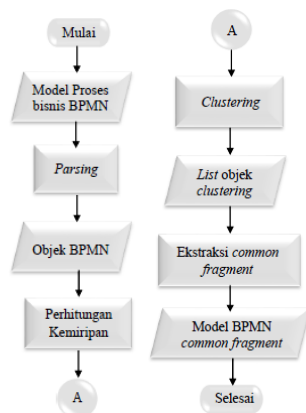
Setelah itu masing-masing nilai kemiripan dikalikan dengan bobot nilai yang telah ditetapkan sesuai dengan kriteria kemiripan masing-masing. Hasil pembobotan penghitungan kemiripan akan digunakan untuk proses *clustering*.

D. Proses *Clustering*

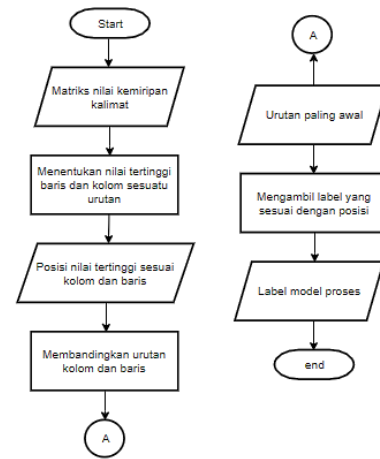
Clustering merupakan pengelompokan objek yang mirip menjadi satu *cluster*. Tujuan dari *graph clustering* adalah untuk mempartisi simpul dalam grafik besar menjadi kelompok yang berbeda berdasarkan berbagai kriteria seperti konektivitas titik atau kesamaan lingkungan. Teknik *graph clustering* sangat berguna untuk mendeteksi kelompok yang terhubung erat dalam grafik besar [9]. Metode yang digunakan pada tahap ini adalah berdasarkan perbandingan nilai *threshold*. Apabila nilai kemiripan bernilai sama dengan atau lebih dari nilai *threshold* maka akan dilakukan ekstraksi *common fragment*, sedangkan nilai kemiripan yang kurang dari nilai *threshold* tidak dilakukan ekstraksi *common fragment*.

E. Ekstraksi *Common Fragment*

Common Fragment secara umum adalah bagian dari proses yang mengandung bagian yang serupa untuk semua variasi proses bisnis [10]. *Common fragment* yang dimaksud disini merupakan rangkaian *node* meliputi *event*, *task* dan *gateway*. *Node* yang dimaksud selalu ada dalam beberapa *graph*. *Common fragment* merupakan *fragment* utama penyusun dari model proses bisnis dalam satu *cluster*. *Common fragment* didapatkan dari penggabungan beberapa *graph* dan menghilangkan *node* yang frekuensinya berkurang. Sehingga proses dalam mendapatkan *common fragment* didapatkan dengan menggabungkan beberapa *graph* menjadi satu *graph*. Dari *graph* gabungan tersebut dilakukan proses *digesting* hingga didapatkan satu *graph* yang merupakan *fragment* utama [11].



Gambar 2. *flowchart* pembentukan *common fragment*

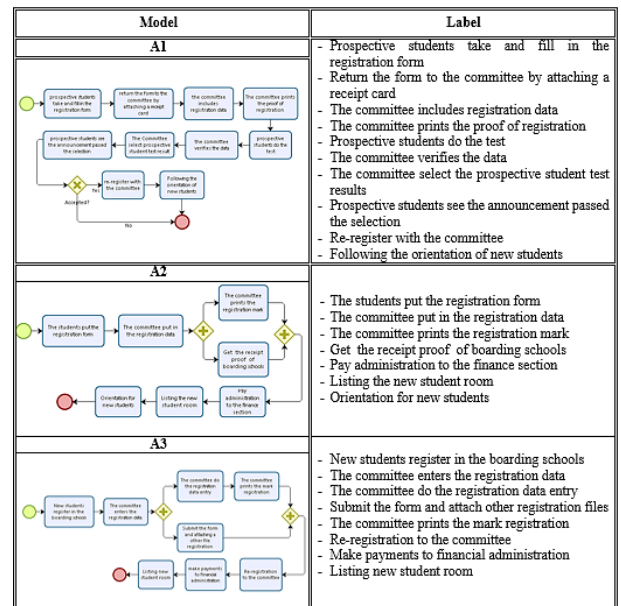


Gambar 3. *flowchart* pembentukan *common fragment* berdasar *contextual similarity*

IV. HASIL DAN PEMBAHASAN

A. *Input Model*

Data yang digunakan dalam penelitian ini berupa model proses bisnis yang terdapat pada ERP Pondok Pesantren. Model proses bisnis yang dijadikan objek dan merupakan sebagai data primer yang berekstensi *.xpd*. Pada ERP Pondok Pesantren, model proses bisnis yang dijadikan sebagai sumber data meliputi proses akademik, kesantrian, dan penerimaan santri baru.



Gambar 4. Model Proses

B. Menghitung Kemiripan

1) *Structural similarity*

Berikut elemen – elemen dari model 1 dan model 2 berdasar strukturalnya:

A1: task1 task2 task3 task4 task5 task6 task7 task8 or1 task9 task10 starttask1 task1task2 task2task3 task3task4 task4task5 task5task6 task6task7 task7task8 task8or1 or1end or1task9 task9task10 task10end

A2: task1 task2 parallel1 task3 task4 task5 parallel2 task6 task7 task8 starttask1 task1task2 task2parallel1 parallel1task3 task3task4 parallel1task5 task4parallel2 task5parallel2 parallel2task6 task6task7 task7end

Berikut perhitungan kemiripan structural menggunakan algoritma jaccard coefficient :

- Pondok Pesantren Al Rifai : 24
- Pondok Pesantren Anwarul Huda : 21
- Jumlah irisan (intersection) : 12

Perhitungan structural similarity :

- Jaccard Coefficient Similarity :

$$J(Model_1, Model_2) = \frac{12}{(24+21)-12} = \frac{12}{33} = 0.36$$

Tabel 2. Matriks kemiripan structural

	A1	A2	A3	B1	B2	B3	C1	C2	C3
A1	1.00	0.33	0.52	0.50	0.46	0.38	0.42	0.41	0.41
A2	0.33	1.00	0.41	0.24	0.27	0.48	0.52	0.41	0.44
A3	0.52	0.41	1.00	0.53	0.34	0.36	0.38	0.36	0.40
B1	0.50	0.24	0.53	1.00	0.28	0.24	0.31	0.30	0.30
B2	0.46	0.27	0.34	0.28	1.00	0.48	0.63	0.59	0.59
B3	0.38	0.48	0.36	0.24	0.48	1.00	0.30	0.28	0.23
C1	0.42	0.52	0.38	0.31	0.63	0.30	1.00	0.86	0.86
C2	0.41	0.41	0.36	0.30	0.59	0.28	0.86	1.00	1.00
C3	0.41	0.44	0.40	0.30	0.59	0.23	0.86	1.00	1.00

2) Behavioral Similarity

Berikut elemen – elemen dari model 1 dan model 2 berdasar behavioral-nya:

A1: task1task2 task1task2task3 task2task3task4 task3task4task5 task4task5task6 task5task6task7 task6task7task8 task7task8 task7task8task9 task8task9task10 task9task10

A2: task1task2 task1task2task3 task1task2task5 task2task3task4 task2task5task6 task3task4task6 task4task6task7 task5task6task7 task6task7task8 task7task8

Berikut perhitungan kemiripan behavioral menggunakan algoritma jaccard coefficient :

- Pondok Pesantren Al Rifai : 11
- Pondok Pesantren Anwarul Huda : 10
- Jumlah irisan (intersection) : 6

Perhitungan behavioral similarity :

- Jaccard Coefficient Similarity :

$$J(Model_1, Model_2) = \frac{6}{(11+10)-6} = \frac{6}{15} = 0,4$$

Tabel 3. Matriks kemiripan behavioral

	A1	A2	A3	B1	B2	B3	C1	C2	C3
A1	1.00	0.25	0.19	0.21	0.31	0.25	0.31	0.29	0.29
A2	0.25	1.00	0.31	0.19	0.19	0.42	0.53	0.41	0.41
A3	0.19	0.31	1.00	0.17	0.40	0.23	0.16	0.15	0.15
B1	0.21	0.19	0.17	1.00	0.10	0.11	0.18	0.17	0.17
B2	0.31	0.19	0.40	0.10	1.00	0.40	0.54	0.50	0.50
B3	0.25	0.42	0.23	0.11	0.40	1.00	0.20	0.19	0.19
C1	0.31	0.53	0.16	0.18	0.54	0.20	1.00	0.79	0.79
C2	0.29	0.41	0.15	0.17	0.50	0.19	0.79	1.00	1.00
C3	0.29	0.41	0.15	0.17	0.50	0.19	0.79	1.00	1.00

3) Semantic Similarity

Menghitung kemiripan antar term:

Contoh kalimat :

- Sentence 1: Return the form to the committee by attaching a receipt card

- Sentence 2 : Submit the form and attaching an other file registration

Dari dua kalimat tersebut kemudian dilakukan perhitungan kemiripan makna antar kata dengan metode *wu palmer* menggunakan WS4J. Di bawah ini adalah hasil matriks perhitungannya :

Tabel 4. Matriks nilai kemiripan kata menggunakan WS4J

	Return /VB	the /DT	form /NN	to /TO	the /DT	committee /NN	by /IN	attaching /VBG	a /DT	receipt /JJ	card /NN
Submit/VB	0.9091	-	-	-	-	-	-	0.4000	-	-	-
the/DT	-	-	-	-	-	-	-	-	-	-	-
form/NN	-	-	1.0000	-	-	0.6667	-	-	-	-	0.6667
and/CC	-	-	-	-	-	-	-	-	-	-	-
attaching/VBG	0.4444	-	-	-	-	-	-	1.0000	-	-	-
a/DT	-	-	-	-	-	-	-	-	-	-	-
other/JJ	-	-	-	-	-	-	-	-	-	-	-
file/NN	-	-	0.6667	-	-	0.5000	-	-	-	-	0.7059
registration/NN	-	-	0.8235	-	-	0.6667	-	-	-	-	0.6250

Tabel 5. Nilai kemiripan tertinggi tiap kriteria

Jenis	Nilai kemiripan kata	Rata - rata
Noun	1, 0.6667, 0.7059	0.79
Verb	0.9091, 1	0.95

Perhitungan Sentence Similarity :

$$Sentence\ similarity = (\bar{n}_{max} \times b_n) + (\bar{v}_{max} \times b_v)$$

$$= (0,79 \times 0,25) + (0,95 \times 0,75)$$

$$= 0,20 + 0,72 = 0,92$$

Berdasarkan perhitungan di atas menunjukkan bahwa nilai sentence similarity bernilai 0,92.

Tabel 6. Perhitungan contextual similarity pada model A1 dan A2

		A2								
		c1	c2	c3	c4	c5	c6	c7	c8	R
A1	b1	0.22	0.67	0.59	0.84	0.54	0.18	0.18	0.65	0.84
	b2	0.18	0.68	0.61	0.91	0.59	0.20	0.62	0.57	0.91
	b3	0.18	0.63	0.63	0.70	0.51	0.21	0.80	0.45	0.80
	b4	0.19	0.53	0.53	0.50	0.98	0.22	0.78	0.42	0.98
	b5	0.22	0.51	0.91	0.54	0.50	0.13	0.87	0.50	0.91
	b6	0.16	0.55	0.55	0.72	0.44	0.19	0.56	0.39	0.72
	b7	0.19	0.44	0.49	0.47	0.45	0.17	0.44	0.42	0.49
	b8	0.18	0.51	0.69	0.73	0.48	0.13	0.78	0.49	0.78
	b9	0.96	0.18	0.24	0.19	0.23	0.98	0.14	0.16	0.98
	b10	0.21	0.73	0.76	0.54	0.70	0.13	0.62	0.52	0.76
Rata – Rata		0.82								

Dari seluruh perbandingan tersebut kemudian setiap perbandingan dicari nilai tertingginya. Pada Tabel 6, kolom R merupakan kolom nilai tertinggi. Kemudian dihitung nilai rata-rata dari nilai tertinggi pada perbandingan tersebut. Sehingga diperoleh nilai sentence similarity secara keseluruhan model proses pada label dalam elemen task sebesar 0,82.

Berikut merupakan hasil nilai kemiripan seluruh model proses:

Tabel 7. Matriks nilai semantic similarity

	A1	A2	A3	B1	B2	B3	C1	C2	C3
A1	1.00	0.77	0.91	0.52	0.45	0.43	0.59	0.51	0.49
A2	0.77	1.00	0.82	0.51	0.31	0.40	0.51	0.61	0.61
A3	0.91	0.82	1.00	0.47	0.27	0.28	0.50	0.44	0.56
B1	0.52	0.51	0.47	1.00	0.70	0.75	0.47	0.42	0.50
B2	0.45	0.31	0.27	0.70	1.00	0.97	0.37	0.40	0.30
B3	0.43	0.40	0.28	0.75	0.97	1.00	0.34	0.40	0.50
C1	0.59	0.51	0.50	0.47	0.37	0.34	1.00	0.93	0.95
C2	0.51	0.61	0.44	0.42	0.40	0.40	0.93	1.00	0.93
C3	0.49	0.61	0.56	0.50	0.30	0.50	0.95	0.93	1.00

C. Pembobotan Kemiripan

Hasil matriks nilai kemiripan *structural, behavioral, dan semantic* diakumulasikan dengan nilai bobot kriteria kemiripan dengan komposisi (0.33, 0.10, 0.57) menghasilkan nilai pada tabel 8 di bawah ini:

Tabel 8. Nilai pembobotan perhitungan kemiripan

	A1	A2	A3	B1	B2	B3	C1	C2	C3
A1	1.00	0.57	0.71	0.48	0.44	0.39	0.51	0.46	0.44
A2	0.57	1.00	0.63	0.39	0.28	0.43	0.52	0.52	0.53
A3	0.71	0.63	1.00	0.46	0.31	0.30	0.42	0.38	0.46
B1	0.48	0.39	0.46	1.00	0.50	0.52	0.39	0.36	0.40
B2	0.44	0.28	0.31	0.50	1.00	0.75	0.47	0.47	0.41
B3	0.39	0.43	0.30	0.52	0.75	1.00	0.31	0.34	0.38
C1	0.51	0.52	0.42	0.39	0.47	0.31	1.00	0.89	0.90
C2	0.46	0.52	0.38	0.36	0.47	0.34	0.89	1.00	0.96
C3	0.44	0.53	0.46	0.40	0.41	0.38	0.90	0.96	1.00

Hasil dari pembobotan penghitungan kemiripan pada (Tabel 8) kemudian digunakan pada proses *clustering*.

D. Melakukan Clustering

Pada proses *clustering* ini dilakukan dengan menentukan nilai *threshold*. Nilai kemiripan yang lebih dari sama dengan *threshold*, maka termasuk dalam *cluster*.

Tabel 9. Hasil clustering dengan nilai threshold (0.49)

A1	A2	A3	B1	B2	B3	C1	C2	C3
1.00	0.57	0.71	0.48	0.44	0.39	0.51	0.46	0.44
0.57	1.00	0.63	0.39	0.28	0.43	0.52	0.52	0.53
0.71	0.63	1.00	0.46	0.31	0.30	0.42	0.38	0.46
0.48	0.39	0.46	1.00	0.50	0.52	0.39	0.36	0.40
0.44	0.28	0.31	0.50	1.00	0.75	0.47	0.47	0.41
0.39	0.43	0.30	0.52	0.75	1.00	0.31	0.34	0.38
0.51	0.52	0.42	0.39	0.47	0.31	1.00	0.89	0.90
0.46	0.52	0.38	0.36	0.47	0.34	0.89	1.00	0.96
0.44	0.53	0.46	0.40	0.41	0.38	0.90	0.96	1.00

Pada Tabel 9 menunjukkan bahwa terdapat tiga kluster. Dari beberapa uji coba yang telah dilakukan diperoleh nilai *threshold* dengan hasil *clustering* yang paling maksimal adalah 0,49. Kluster 1 adalah kelompok proses penerimaan santri baru, kluster 2 adalah kelompok proses kesantrian, dan kluster 3 adalah kelompok proses akademik.

E. Menghitung Akurasi

Setelah semua kombinasi diujicobakan pada setiap algoritma kemiripan, langkah selanjutnya menghitung akurasi dari setiap hasil proses *clustering* yang telah dilakukan. Tujuan dari proses perhitungan akurasi dimaksudkan untuk mengetahui kombinasi yang tepat, algoritma apa yang digunakan, dan nilai *threshold* berapa sehingga dapat menemukan *common fragment* yang akurat. Metode yang digunakan untuk mengukur akurasi adalah ROC. Hasil akurasi didapat dengan rumus :

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (3)$$

TP: True Positive TN: True Negative
FP: False Positive FN: False Negative

Akurasi dari hasil *clustering* dengan komposisi pembobotan (0.33, 0.10, 0.57) dengan nilai *threshold* 0.49 :

$$Accuracy = \frac{27+46}{27+46+8+0} \times 100\%$$

Accuracy = 90,12%

F. Ekstraksi Common Fragment

Setelah mendapatkan hasil *cluster* pada proses *clustering*, kemudian dilakukan proses *common fragment extraction* untuk memperoleh hasil model proses bisnis yang baru. *Common fragment* merupakan *fragment* utama penyusun dari model proses bisnis dalam satu *cluster*. Untuk mendapatkan *fragment* dari model proses dengan melakukan *intersection* dari setiap aktivitas.

Di bawah ini merupakan proses *common fragment extraction*:

Contextual Similarity

Tabel 10. Nilai kemiripan contextual (Common Fragment)

		A2								
		c1	c2	c3	c4	c5	c6	c7	c8	R
A1	b1	0.22	0.67	0.59	0.84	0.54	0.18	0.18	0.65	0.84
	b2	0.18	0.68	0.61	0.91	0.59	0.20	0.62	0.57	0.91
	b3	0.18	0.63	0.63	0.70	0.51	0.21	0.80	0.45	0.80
	b4	0.19	0.53	0.53	0.50	0.98	0.22	0.78	0.42	0.98
	b5	0.22	0.51	0.91	0.54	0.50	0.13	0.87	0.50	0.91
	b6	0.16	0.55	0.55	0.72	0.44	0.19	0.56	0.39	0.72
	b7	0.19	0.44	0.49	0.47	0.45	0.17	0.44	0.42	0.49
	b8	0.18	0.51	0.69	0.73	0.48	0.13	0.78	0.49	0.78
	b9	0.96	0.18	0.24	0.19	0.23	0.98	0.14	0.16	0.98
	b10	0.21	0.73	0.76	0.54	0.70	0.13	0.62	0.52	0.76

Tabel 11. Label Model A1 dan Model A2 (Common Fragment)

A1	A2
1. Prospective students take and fill in the registration form	1. New students register in the boarding schools
2. Return the form to the committee by attaching a receipt card	2. The committee enters the registration data
3. The committee includes registration data	3. The committee do the registration data entry
4. The committee prints the proof of registration	4. Submit the form and attach other registration files
5. Prospective students do the test	5. The committee prints the mark registration
6. The committee verifies the data	6. Re-registration to the committee
7. The committee select the prospective student test results	7. Make payments to financial administration
8. Prospective students see the announcement passed the selection	8. Listing new student room
9. Re-register with the committee	
10. Following the orientation of new students	

Structural Similarity:

- task1 task2 task3 task4 task5 task6 task7 task8 or1 task9 task10 starttask1 task1task2 task2task3 task3task4 task4task5 task5task6 task6task7 task7task8 task8or1 or1end or1task9 task9task10 task10end
- task1 task2 paralle11 task3 task4 task5 paralle12 task6 task7 task8 starttask1 task1task2

task2parallel1 parallel1task3 parallel1task5
 task3task4 task5parallel2 task4parallel2
 parallel2task6 task6task7 task7task8 task8end

intersection:

- task1 task2 task3 task4 task5 task6 task7 task8 starttask1 task1task2 task3task4 task6task7 task7task8

Behavioral Similarity

- task1task2 task1task2task3 task2task3task4 task3task4task5 task4task5task6 task5task6task7 task6task7task8 task7task8 task7task8task9 task8task9task10 task9task10
- task1task2 task1task2task3 task1task2task5 task2task3task4 task2task5task6 task3task4task6 task4task6task7 task5task6task7 task6task7task8 task7task8

intersection:

- task1task2 task1task2task3 task2task3task4 task5task6task7 task6task7task8 task7task8

Common Fragment :

Semantic

1. Prospective students take and fill in the registration form
2. Return the form to the committee by attaching a receipt card
3. The committee do the registration data entry
4. The committee prints the proof of registration
5. Prospective students do the test
6. Re-registration to the committee
7. Make payments to financial administration
8. Prospective students see the announcement passed the selection
9. Re-register with the committee
10. Following the orientation of new students

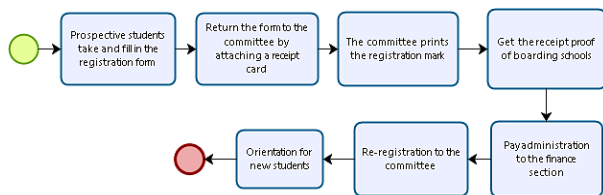
Structural

- task1 task2 task3 task4 task5 task6 task7 task8 starttask1 task1task2 task3task4 task6task7 task7task8

Behavioral

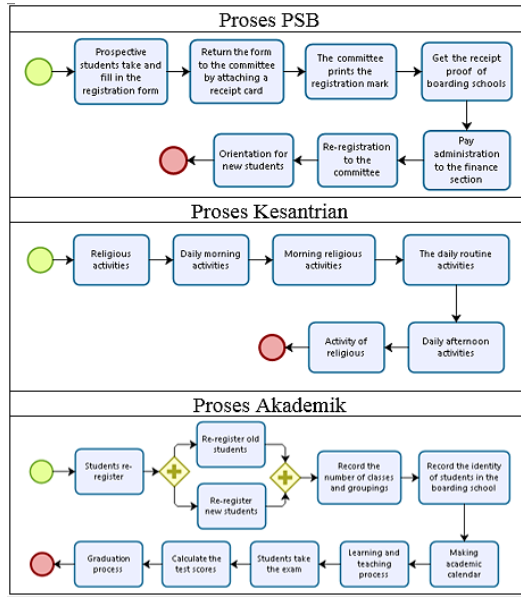
- task1task2 task1task2task3 task2task3task4 task5task6task7 task6task7task8 task7task8

Untuk menentukan model BPMN yaitu dengan menyesuaikan antara *common fragment* dari *label* dan struktur *element task*. Berikut hasil ekstraksi *common fragment* model BPMN:



Gambar 5. Hasil Model Proses Ekstraksi *Common Fragment*

Berikut adalah hasil *Common Fragment* dari semua proses:



Gambar 6. Hasil ekstraksi *common fragment*

Pada Gambar 6 merupakan hasil ekstraksi *common fragment* dari proses penerimaan santri baru, proses kesantrian, dan proses akademik.

V. KESIMPULAN

Komposisi nilai bobot dari kriteria kelas kata yang diberikan untuk perhitungan *sentence similarity* adalah 0,75 untuk kelas *verb*, dan 0,25 untuk kelas *noun*. Komposisi nilai bobot kemiripan adalah 0,10 untuk *behavioral*, 0,33 untuk *structural*, dan 0,57 untuk *semantic*. Proses *clustering* dilakukan dengan membandingkan nilai akumulasi perhitungan pembobotan kemiripan dengan nilai *threshold* yang telah ditentukan. Apabila nilai kemiripan (\geq) nilai *threshold* maka nilai tersebut lolos *cluster*. Nilai *threshold* sebesar 0,49.

REFERENSI

[1] M. Dumas, L. Garcia-Banuelos and R. Dijkman, "Similarity Search of Business Process Models," *IEEE Bulletin of the Technical Committee on Data Engineering*, vol. 32, no. 3, pp. 23-28, 2009.

[2] A. C. Fauzan, R. Sarno, M. A. Yaqin and A. Jamal, "Extracting common fragment based on behavioral similarity using transition adjacency relations for scalable business processes," *International Conference on Information & Communication Technology and System (ICTS)*, 2017.

[3] A. W. Prasetya, "Aplikasi Manajemen Proses Bisnis Menggunakan Metode Analisis Kemiripan Behavioral, Structural, dan Semantic untuk Meningkatkan Akurasi dalam Penentuan Common Fragment Workflow pada ERP Pesantren," Universitas Islam Negeri Maulana Malik Ibrahim Malang, Malang, 2017.

- [4] A. Maulana, M. Bijaksana Arif and M. Syahrul Mubarak, "Perancangan Semantic Similarity based on Word Thesaurus Menggunakan Pengukuran Omiotis Untuk Pencarian Aplikasi pada I-GRACIAS," *e-Proceeding of Engineering*, vol. 3, no. 2, p. 3689, 2016.
- [5] A. W. Prasetya, M. A. Yaqin and Syahiduzzaman, "Common Process Extraction Pada Scalable Model Proses Bisnis," *Konferensi Nasional Sistem Informasi*, pp. 8-9, 2018.
- [6] R. Dijkman, M. Dumas and L. Garcia-Banuelos, "Graph Matching Algorithms for Business Process Model Similarity Search," *Proceeding BPM Ulm, Germany*, 2009.
- [7] J. Mendling, B. v. Dongen and W. v. d. Aalast, "On the Degree of Behavioral Similarity between Business," *Workshop der Gesellschaft für Informatik e.V. (GI) und Treffen ihres Arbeitskreises "Geschäftsprozessmanagement mit Ereignisgesteuerten Prozessketten (WI-EPK)*, vol. 303, pp. 39-58, 2007.
- [8] L. Y. Banowosari and I. W. I. Wicaksana, "Pengembangan Aplikasi Antar-muka Pemakai untuk Penghitungan Similaritas Semantik Berbasis String dan Wordnet," vol. 1, no. 2, pp. 10-12, 2007.
- [9] Y. Zhou, H. Cheng and J. Xu Yu, "Graph Clustering Based on Structural/ Attribute Similarities," *Proceedings of the VLDB Endowment*, vol. 2, no. 1, pp. 718-729, 2009.
- [10] M. L. Rosa, M. Dumas, U. Reina and R. Dijkman, "Business Process Model Merging: An Approach to Business Process Consolidation," *ACM Transactions on Software Engineering and Methodology*, vol. 22, no. 2, pp. 1-42, 2013.
- [11] R. Sarno, E. W. Pamungkas, D. Sunaryono and Sarwosri, "Workflow Common Fragments Extraction Based on WSDL Similarity and Graph Dependency," *IEEE International Seminar on Intelligent Technology and Its Applications*, pp. 309-314, 2015.