

Analisis Kemiripan Model Proses Bisnis PMBoK dan Scrum menggunakan Metode *Jaccard Coefficient Similarity* dan *Semantic Similarity*

Ibrohim M. Aqimuddin^{1,*}, Seta M. Pamungkas², Chandra Gunawan³, Muhammad Ainul Yaqin⁴, Abd. Charis Fauzan⁵

^{1,2,3,4}Jurusan Teknik Informatika, Universitas Islam Negeri Maulana Malik Ibrahim, Indonesia

¹Prodi Ilmu Komputer, Universitas Nahdlatul Ulama Blitar, Indonesia

¹18650067@student.uin-malang.ac.id; ²18650043@student.uin-malang.ac.id; ³18650071@student.uin-malang.ac.id;

⁴yaqinov@ti.uin-malang.ac.id, ⁵abdcharis@unublitar.ac.id

*corresponding author

INFO ARTIKEL

Sejarah Artikel

Diterima: 23 Desember 2021
Direvisi: 30 April 2023
Diterbitkan: 30 Agustus 2023

Kata Kunci

Model Proses Bisnis
PMBoK
Scrum
Semantic Similarity
Behavioral Similarity

ABSTRAK

Pengukuran perbandingan kemiripan model bisnis PMBoK dan Scrum dilakukan untuk mengetahui titik nilai tertinggi antara kedua objek perbandingan. Fokus perbandingan menggunakan model proses bisnis PMBoK dan Scrum. Kedua model proses bisnis tersebut memiliki sifat yang berbeda sehingga dengan diketahuinya titik nilai tertinggi dan titik nilai terendah persamaan keduanya, diharapkan dapat digunakan untuk pendukung dalam pembangunan aplikasi manajemen proyek. Perhitungan ini menggunakan perbandingan kemiripan *semantic* menggunakan Algoritma Wu Palmer, *structural* dan *behavioral* yang keduanya menggunakan Algoritma *Jaccard Coefficient Similarity*. Ketiga perbandingan itu lalu dicari nilai akumulasi menggunakan pembobotan dengan nilai bobot yang berbeda setiap kemiripan. Hasil akumulasi tersebut lalu dikelompokkan/*clustering* dengan batas nilai minimal. Hasil pengelompokkan nilai lalu dicari tingkat akurasi untuk mengetahui seberapa besar akurasi yang didapat. Hasil perhitungan nilai pembobotan didapatkan komposisi nilai bobot *semantic* sebesar 0.57, *structural* sebesar 0.33, dan *behavioral* sebesar 0.1. Pada proses pengelompokkan/*clustering* menggunakan nilai batas minimal yaitu 0.49 dan nilai batas maksimal yaitu 0.55. Sehingga dari perhitungan tersebut didapatkan sebuah nilai persamaan tertinggi dan terendah.

PENDAHULUAN

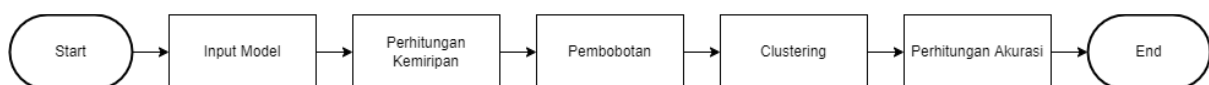
Pengembangan sebuah *software* tidak selalu berjalan dengan mulus. Tingkat kegagalan dalam pengembangan dapat dikatakan lebih besar daripada tingkat keberhasilan [1]. Sehingga untuk menekan tingkat kegagalan sebuah proyek, dibutuhkan sebuah model proses manajemen proyek agar pengerjaan bersifat lebih teratur dan terarah. Setiap pengembang memiliki metode yang berbeda dalam pengembangan mereka dalam mencapai target yang telah ditentukan [2]. Sehingga hal ini memicu sebuah fokus tersendiri pada pengembangan model proses manajemen proyek yang pada akhirnya menyebabkan bermunculannya model proses yang bermacam-macam dan saling bervariasi [3]. Setiap model proses memiliki kelebihan dan kekurangannya sendiri-sendiri. Banyaknya varian model proses bisnis tersebut menyebabkan para pengembang kebingungan dalam menentukan model proses bisnis yang akan digunakan. Kebingungan tersebut memicu pengembangan sebuah aplikasi manajemen proyek yang dapat membantu para pengguna, akan tetapi aplikasi manajemen proyek memerlukan nilai analisis kemiripan sebagai dasar dalam penentuan hasil akhir.

Sehingga penelitian pada paper ini bertujuan untuk menganalisis persamaan 2 model proses bisnis yang sering digunakan [4]. Model proses bisnis yang akan digunakan adalah Model Proses Bisnis PMBOK yang terfokus pada perencanaan proyek yang matang di awal lalu dieksekusi sesuai dengan perencanaan yang telah dikembangkan [5], sedangkan Model Proses Bisnis Scrum yang terfokus pada pengerjaan proyek besar yang dibagi menjadi bagian-bagian kecil lalu dikerjakan secara cepat dan adaptif [6]. Kedua proses bisnis tersebut memiliki kelebihan dan kekurangan dalam setiap struktur dan aktivitas di dalamnya sehingga perhitungan ini menggunakan acuan awal yaitu proses bisnis PMBoK dengan 10 *knowledges*-nya. Model Proses Bisnis Scrum akan dianalisis setiap aktivitasnya dan diklasifikasikan menjadi *knowledges*-nya berdasarkan *knowledges* PMBoK, lalu dibandingkan persamaannya antara satu dan lainnya untuk mengetahui bagian struktur proses bisnis yang memiliki sebuah nilai titik persamaan tertinggi dan nilai titik persamaan terendah.

Kedua model tersebut akan direpresentasikan terlebih dahulu berupa BPMN (*Business Process Modelling and Notation*) sebelum dilakukan perhitungan. Perhitungan kemiripan kedua model ini akan dihitung dengan metode *semantic similarity* yaitu metode perhitungan setiap kalimat antara kedua model dan fokus perhitungan metode ini akan diwakilkan menggunakan Algoritma Wu Palmer. Perhitungan kemiripan semantik akan menggunakan *tools* WS4J untuk mendapatkan nilai yang akurat. Selain itu, setiap model proses bisnis juga akan dihitung nilai kemiripan dari struktur dan aktivitas di dalamnya menggunakan Metode *Jaccard Coefficient Similarity*. Dari perbandingan ini akan dihasilkan 3 matriks besar yang berisikan nilai perbandingan dari kemiripan semantik, kemiripan struktur, dan kemiripan aktivitas.

Hasil perbandingan 3 kemiripan tersebut lalu digabungkan dengan pemberian bobot setiap jenis kemiripan sehingga mendapatkan nilai gabungan dari nilai 3 persamaan tersebut. Nilai tersebut ditampung dalam matrik besar untuk mempermudah proses selanjutnya. Matrik yang berisi nilai persamaan itu lalu diklasifikasikan atau *clustering* dengan nilai minimal 0.55. Apabila nilai perbandingan melebihi nilai minimal, maka perbandingan tersebut lolos dari proses *clustering*. Sehingga nilai pada matriks akan terklasifikasikan untuk perhitungan nilai akurasi dalam penentuan kombinasi. Hasil dari perbandingan ini diharapkan dapat digunakan sebagai langkah awal dalam pembuatan aplikasi manajemen proyek.

METODE



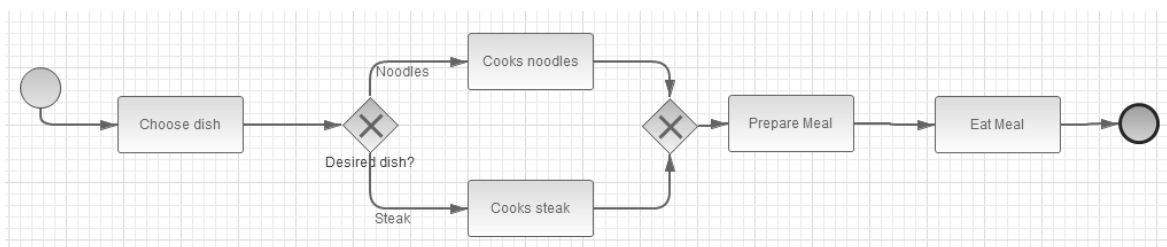
Gambar 1. Alur Penelitian

Gambar 1 merepresentasikan setiap tahapan dalam penelitian yang dimulai dari *Input Model* kedua model proses bisnis hingga perhitungan akurasi dari kedua nilai kemiripan yang telah dihitung. Tahapan ini akan dilakukan dengan tujuan untuk mencari nilai similaritas antara model proses bisnis PMBoK dan Scrum.

Input Model

Data yang dipakai untuk melakukan perhitungan yaitu berupa file BPMN dengan ekstensi *xpdl*. *Business Process Modelling Notation* (BPMN) merupakan suatu pendekatan manajemen holistik untuk meningkatkan efektivitas dan efisiensi bisnis seiring upaya untuk mencapai inovasi, fleksibilitas dan integrasi dengan teknologi [5].

Setiap bagian dalam BPMN mendefinisikan setiap proses dan alur sebuah proses bisnis. Oleh karena itu, BPMN dapat digunakan untuk sebagai data untuk dilakukan sebuah perhitungan. Gambar 2 merepresentasikan gambaran umum dari BPMN.



Gambar 2. Contoh BPMN

Menghitung Kemiripan

Kemiripan BPMN dikaji secara kemiripan struktural (*Structural Similarity*), kemiripan perilaku (*Behavioral Similarity*) dan kemiripan semantik (*Semantic Similarity*). Sebab BPMN PMBoK ataupun Scrum memiliki pola-pola dan kemiripan walaupun keduanya merupakan model proses bisnis yang berbeda. Perhitungan kemiripan merupakan solusi pergantian *framework* bisnis sehingga organisasi tidak perlu mengulang pembuatan proses bisnis yang baru cukup dengan melakukan pengecekan kemiripan proses bisnis sehingga mereka hanya perlu *redesign* proses bisnis sesuai dengan hasil kesamaan [7].

Structural Similarity

Kesamaan struktural merupakan nilai kesamaan antara model proses bisnis berdasarkan bentuknya. Dengan memetakan bentuk kedua model proses bisnis, metode Jaccard *Similarity* digunakan untuk menghitung kemiripan tersebut. Jaccard *Similarity* atau juga dikenal dengan koefisien Jaccard merupakan sebuah metode statistika untuk mencari kesamaan dan keberagaman antara dua set terbatas. Jaccard merupakan metode yang sederhana dan mudah digunakan [8]. Pada persamaan (1), koefisien Jaccard didapat dari perhitungan irisan himpunan *A* dengan himpunan *B* dibagi dengan gabungan himpunan *A* dengan himpunan *B*.

$$SimJ(A,B) = \frac{|A \cup B|}{|A \cap B|} \quad (1)$$

Behavioral Similarity

Behavioral similarity merupakan teknik perhitungan derajat kemiripan proses bisnis berdasarkan kesamaan perilaku antara proses bisnis yang dibandingkan. Untuk mengukur nilai kesamaan perilaku, digunakan metode Causal Footprints. Causal Footprints digunakan untuk memperoleh kumpulan kondisi pada urutan aktifitas - aktifitas yang memegang setiap kasus model proses dan memanfaatkannya untuk *reasoning* sebuah *correctness* [7].

Semantic Similarity

Semantic similarity mencari tingkatan kesetaraan semantik antara dua hal yang dapat berupa konsep, kalimat atau dokumen [9]. Metode statistika yang digunakan untuk menghitung kemiripan semantik adalah metode Wu & Palmer dan Analytical Hierarchy Process (AHP). Perhitungan Wu & Palmer dilakukan berdasarkan kemiripan makna antar kata pada kalimat [10]. Lalu, menentukan kemiripan kalimat dengan mengakumulasi matriks nilai kemiripan kata dengan bobot nilai kelas kata dengan menggunakan metode AHP. Kriteria yang digunakan adalah kata benda (noun) dan kata kerja (verb) [12].

Persamaan (2) merupakan persamaan untuk menghitung sentence *similarity*. n_{max} merupakan nilai rata-rata kriteria noun (kata kerja) dan v_{max} merupakan nilai rata-rata dari kriteria verb (kata kerja). Kemudian b adalah sebuah bobot dari masing-masing input. Kata kerja lebih sedikit ditemukan pada objek yang akan dibandingkan daripada kata benda sehingga ditetapkan bobot keduanya yaitu 0,75 untuk kata kerja dan 0,25 untuk kata benda.

$$\text{Semantic Similarity} = (\underline{n_{max}} \times b_n) + (\underline{v_{max}} \times b_v) \quad (2)$$

Pembobotan

Pembobotan merupakan nilai *weight* (berat) masing-masing kriteria kemiripan yaitu struktural, perilaku dan semantik. Untuk menentukan bobot setiap kriteria, dilakukan dengan cara menjumlahkan hasil kali nilai kemiripan dengan masing-masing bobot kriteria kemiripan yang telah ditentukan. Metode yang dipakai untuk menghitung bobot adalah metode AHP (*Analytical Hierarchy Process*). AHP adalah metode untuk mengambil keputusan berdasarkan campuran beberapa kriteria kuantitatif maupun kualitatif yang diformulasikan secara hirarkis [11].

Tabel 1. Nilai bobot kriteria

Kriteria	Nilai bobot prioritas
Struktural	0.33
<i>Behavioral</i>	0.10
Semantik	0.57

Proses Clustering

Clustering adalah pengelompokan data atau objek berdasarkan kemiripan satu sama lain. Teknik yang digunakan untuk melakukan *clustering* adalah *graph clustering*. *Graph clustering* melihat pola di antara objek-objek pada sebuah graf untuk memproduksi sebuah pengelompokan objek yang berarti [12]. *Clustering* dilakukan untuk mengeliminasi kelompok-kelompok objek yang memiliki nilai kemiripan yang rendah. Sehingga hanya kelompok-kelompok objek yang memiliki nilai kemiripan melebihi atau sama dengan nilai *threshold* yang akan dilakukan proses ekstraksi *common fragment* [13].

Menghitung Akurasi

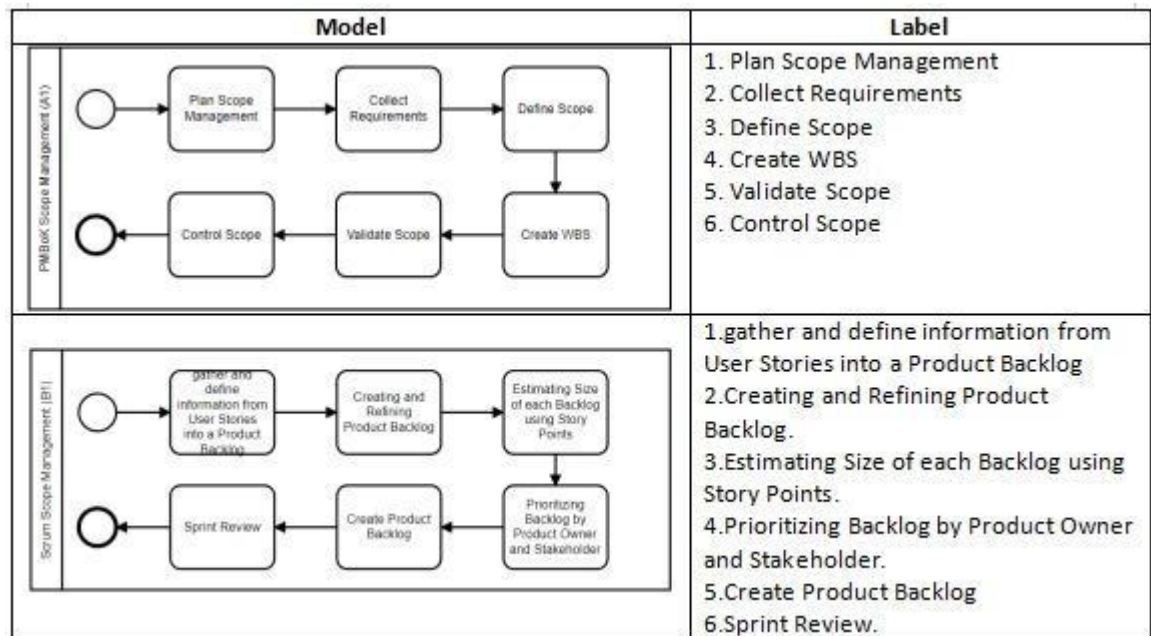
Setelah *clustering*, perhitungan akurasi dilakukan. Akurasi bertujuan mengevaluasi perhitungan atau model yang dihasilkan. Metode *Receiver Operating Characteristic* (ROC) dipakai untuk menghitung akurasi. ROC adalah sebuah teknik untuk memvisualkan, mengorganisir dan memilih penggolongan berdasarkan performanya. Persamaan (3) merupakan persamaan untuk menghitung akurasi. TP, TN, FP, FN merupakan berturut-turut jumlah benar positif (*true positive*), jumlah benar negatif (*true negative*), jumlah salah positif (*false positive*), dan jumlah salah negatif (*false negative*).

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

HASIL DAN PEMBAHASAN

Input Model

Data input yang digunakan pada penelitian ini adalah proses bisnis yang diubah terlebih dahulu ke dalam bentuk BPMN sebagai langkah awal perbandingan setiap event. Knowledge model bisnis PMBoK diambil referensi dari buku *A Guide to the Project Management Body of Knowledge*. Sedangkan untuk model bisnis Scrum, setiap fungsinya akan dipilah sehingga dapat diklasifikasikan pada setiap knowledge seperti proses model bisnis PMBoK. Gambar 3 adalah salah satu knowledge kedua proses bisnis yang telah diubah menjadi format BPMN.



Gambar 3. BPMN PMBoK (atas) dan BPMN Scrum (bawah)

Perhitungan kemiripan ini tentunya akan menggunakan beberapa parameter sehingga pada Gambar 4 ditampilkan parameter-parameter yang akan digunakan.

Parameter	Knowledge Areas
A1	PMBok Scope Management
A2	PMBok Schedule Management
A3	PMBok Cost Management
A4	PMBok Quality Management
A5	PMBok Resource Management
A6	PMBok Communication Management
A7	PMBok Risk Management
A8	PMBok Procurement Management
A9	PMBok Stakeholder Management
A10	PMBok Integration Management
B1	Scrum Scope Management
B2	Scrum Schedule Management
B3	Scrum Cost Management
B4	Scrum Quality Management
B5	Scrum Resource Management
B6	Scrum Communication Management
B7	Scrum Risk Management
B8	Scrum Procurement Management
B9	Scrum Stakeholder Management
B10	Scrum Integration Management

Gambar 4. Knowledge Area (task) PMBoK dan Scrum

Menghitung Kemiripan

1) *Structural Similarity*

Berikut elemen-elemen dari model A1 dan model B1 berdasarkan strukturnya :

A1 : task1 task2 task3 task4 task5 task6 starttask1 task1task2 task2task3 task3task4 task4task5 task5task6 task6end

B1 : task1 task2 task3 task4 task5 task6 starttask1 task1task2 task2task3 task3task4 task4task5 task5task6 task6end

Berikut perhitungan kemiripan *structural* menggunakan algoritma jaccard coefficient :

- PMBoK Business Scope Management : 13
- Scrum Business Scope Management : 13
- Jumlah irisan (intersection) : 13

Perhitungan *structural similarity* :

- *Jaccard Coefficient Similarity* :

$$J(A1;B1) = \frac{13}{(13+13)-13} = \frac{13}{13} = 1$$

Setiap hasil perbandingan lalu dimasukkan ke dalam matriks besar pada gambar 5 di bawah ini.

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10
A1	1.00	1.00	0.69	0.54	1.00	0.54	1.00	0.54	0.69	0.87	0.85	0.85	0.38	0.38	0.85	0.87	0.00	0.00	0.54	0.87
A2	1.00	1.00	0.69	0.54	1.00	0.54	1.00	0.54	0.69	0.87	0.85	0.85	0.38	0.38	0.85	0.87	0.00	0.00	0.54	0.87
A3	0.69	0.69	1.00	0.78	0.69	0.78	0.69	0.78	1.00	0.60	0.82	0.82	0.56	0.56	0.82	0.60	0.00	0.00	0.78	0.60
A4	0.54	0.54	0.78	1.00	0.54	1.00	0.54	1.00	0.78	0.47	0.64	0.64	0.71	0.71	0.64	0.47	0.00	0.00	1.00	0.47
A5	1.00	1.00	0.69	0.54	1.00	0.54	1.00	0.54	0.69	0.87	0.85	0.85	0.38	0.38	0.85	0.87	0.00	0.00	0.54	0.87
A6	0.54	0.54	0.78	1.00	0.54	1.00	0.54	1.00	0.78	0.47	0.64	0.64	0.71	0.71	0.64	0.47	0.00	0.00	1.00	0.47
A7	1.00	1.00	0.69	0.54	1.00	0.54	1.00	0.54	0.69	0.87	0.85	0.85	0.38	0.38	0.85	0.87	0.00	0.00	0.54	0.87
A8	0.54	0.54	0.78	1.00	0.54	1.00	0.54	1.00	0.78	0.47	0.64	0.64	0.71	0.71	0.64	0.47	0.00	0.00	1.00	0.47
A9	0.69	0.69	1.00	0.78	0.69	0.78	0.69	0.78	1.00	0.60	0.82	0.82	0.56	0.56	0.82	0.60	0.00	0.00	0.78	0.60
A10	0.87	0.87	0.60	0.47	0.87	0.47	0.87	0.47	0.60	1.00	0.73	0.73	0.33	0.33	0.73	1.00	0.00	0.00	0.47	1.00
B1	0.85	0.85	0.82	0.64	0.85	0.64	0.85	0.64	0.82	0.73	1.00	1.00	0.45	0.45	1.00	0.73	0.00	0.00	0.64	0.73
B2	0.85	0.85	0.82	0.64	0.85	0.64	0.85	0.64	0.82	0.73	1.00	1.00	0.45	0.45	1.00	0.73	0.00	0.00	0.64	0.73
B3	0.38	0.38	0.56	0.71	0.38	0.71	0.38	0.71	0.56	0.33	0.45	0.45	1.00	1.00	0.45	0.33	0.00	0.00	0.71	0.33
B4	0.38	0.38	0.56	0.71	0.38	0.71	0.38	0.71	0.56	0.33	0.45	0.45	1.00	1.00	0.45	0.33	0.00	0.00	0.71	0.33
B5	0.85	0.85	0.82	0.64	0.85	0.64	0.85	0.64	0.82	0.73	1.00	1.00	0.45	0.45	1.00	0.73	0.00	0.00	0.64	0.73
B6	0.87	0.87	0.60	0.47	0.87	0.47	0.87	0.47	0.60	1.00	0.73	0.73	0.33	0.33	0.73	1.00	0.00	0.00	0.47	1.00
B7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-	-	0.00	0.00
B8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-	-	0.00	0.00
B9	0.54	0.54	0.78	1.00	0.54	1.00	0.54	1.00	0.78	0.47	0.64	0.64	0.71	0.71	0.64	0.47	0.00	0.00	1.00	0.47
B10	0.87	0.87	0.60	0.47	0.87	0.47	0.87	0.47	0.60	1.00	0.73	0.73	0.33	0.33	0.73	1.00	0.00	0.00	0.47	1.00

Gambar 5. Matriks kemiripan *structural*

2) Behavioral Similarity

Berikut elemen-elemen dari model A1 dan model B1 berdasarkan strukturnya :

A1 : task1task2 task1task2task3 task2task3task4 task3task4task5
task4task5task6 task5task6

B1 : task1task2 task1task2task3 task2task3task4 task3task4task5
task4task5task6 task5task6

Berikut perhitungan kemiripan *behavioral* menggunakan algoritma jaccard coefficient :

- PMBoK Business Scope Management : 6
- Scrum Business Scope Management : 6
- Jumlah irisan (intersection) : 6

Perhitungan *behavioral similarity* :

- *Jaccard Coefficient Similarity* :

$$J(A1;B1) = \frac{6}{(6+6)-6} = \frac{6}{6} = 1$$

Setiap perbandingan lalu dimasukkan ke dalam matriks besar seperti pada gambar 6.

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10
A1	1.00	1.00	0.67	0.50	1.00	0.50	1.00	0.50	0.67	0.86	0.83	0.83	0.33	0.33	0.83	0.86	0.00	0.00	0.50	0.86
A2	1.00	1.00	0.67	0.50	1.00	0.50	1.00	0.50	0.67	0.86	0.83	0.83	0.33	0.33	0.83	0.86	0.00	0.00	0.50	0.86
A3	0.67	0.67	1.00	0.75	0.67	0.75	0.67	0.75	1.00	0.57	0.80	0.80	0.50	0.50	0.80	0.57	0.00	0.00	0.75	0.57
A4	0.50	0.50	0.75	1.00	0.50	1.00	0.50	1.00	0.75	0.43	0.60	0.60	0.67	0.67	0.60	0.43	0.00	0.00	1.00	0.43
A5	1.00	1.00	0.67	0.50	1.00	0.50	1.00	0.50	0.67	0.86	0.83	0.83	0.33	0.33	0.83	0.86	0.00	0.00	0.50	0.86
A6	0.50	0.50	0.75	1.00	0.50	1.00	0.50	1.00	0.75	0.43	0.60	0.60	0.67	0.67	0.60	0.43	0.00	0.00	1.00	0.43
A7	1.00	1.00	0.67	0.50	1.00	0.50	1.00	0.50	0.67	0.86	0.83	0.83	0.33	0.33	0.83	0.86	0.00	0.00	0.50	0.86
A8	0.50	0.50	0.75	1.00	0.50	1.00	0.50	1.00	0.75	0.43	0.60	0.60	0.67	0.67	0.60	0.43	0.00	0.00	1.00	0.43
A9	0.67	0.67	1.00	0.75	0.67	0.75	0.67	0.75	1.00	0.57	0.80	0.80	0.50	0.50	0.80	0.57	0.00	0.00	0.75	0.57
A10	0.86	0.86	0.57	0.43	0.86	0.43	0.86	0.43	0.57	1.00	0.71	0.71	0.29	0.29	0.71	1.00	0.00	0.00	0.43	1.00
B1	0.83	0.83	0.80	0.60	0.83	0.60	0.83	0.60	0.80	0.71	1.00	1.00	0.40	0.40	1.00	0.71	0.00	0.00	0.60	0.71
B2	0.83	0.83	0.80	0.60	0.83	0.60	0.83	0.60	0.80	0.71	1.00	1.00	0.40	0.40	1.00	0.71	0.00	0.00	0.60	0.71
B3	0.33	0.33	0.50	0.67	0.33	0.67	0.33	0.67	0.50	0.29	0.40	0.40	1.00	1.00	0.40	0.29	0.00	0.00	0.67	0.29
B4	0.33	0.33	0.50	0.67	0.33	0.67	0.33	0.67	0.50	0.29	0.40	0.40	1.00	1.00	0.40	0.29	0.00	0.00	0.67	0.29
B5	0.83	0.83	0.80	0.60	0.83	0.60	0.83	0.60	0.80	0.71	1.00	1.00	0.40	0.40	1.00	0.71	0.00	0.00	0.60	0.71
B6	0.86	0.86	0.57	0.43	0.86	0.43	0.86	0.43	0.57	1.00	0.71	0.71	0.29	0.29	0.71	1.00	0.00	0.00	0.43	1.00
B7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-	-	0.00	0.00
B8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-	-	0.00	0.00
B9	0.50	0.50	0.75	1.00	0.50	1.00	0.50	1.00	0.75	0.43	0.60	0.60	0.67	0.67	0.60	0.43	0.00	0.00	1.00	0.43
B10	0.86	0.86	0.57	0.43	0.86	0.43	0.86	0.43	0.57	1.00	0.71	0.71	0.29	0.29	0.71	1.00	0.00	0.00	0.43	1.00

Gambar 6. Matriks kemiripan *behavioral*3) *Semantic Similarity*

Menghitung kemiripan antar kalimat dalam setiap task.

Contoh kalimat :

- Kalimat 1 : *Planning Scope Management*
- Kalimat 2 : *gathering and defining information from user stories into a product backlog*

Kedua kalimat tersebut dibandingkan menggunakan WS4J dengan algoritma pembandingan Wu Palmer. Setelah menginputkan kalimat ke dalam WS4J, muncul sebuah matrik perbandingan setiap kata yang diperlihatkan pada Gambar 7.

	planning /VBG	Scope /NNP	Management /NNP
gathering/NN	-	0.6000	0.8235
and/CC	-	-	-
defining/VBG	0.3333	-	-
information/NN	-	0.4286	0.5714
from/IN	-	-	-
user/NN	-	0.5263	0.2857
stories/NNS	-	0.6667	0.4286
into/IN	-	-	-
a/DT	-	-	-
product/NN	-	0.6667	0.5333
backlog/NN	-	0.4286	0.5714

Gambar 7. Hasil perhitungan algoritma Wu Palmer

Sehingga rata-rata dapat dihitung menjadi :

Tabel 2. Hasil perhitungan rata-rata *Noun* dan *Verb*

Type	Nilai	Rata-rata
Noun (Kata Benda)	0.6, 0.823, 0.428, 0.571, 0.52, 0.28, 0.67, 0.42, 0.67, 0.53, 0.42, 0.57	0.542
Verb (Kata Kerja)	0.333	0.333

Dengan menggunakan rata-rata ini, didapatkan nilai sentence similarity sebagai berikut :

$$Similarity = (\underline{n_{max}} \times b_n) + (\underline{v_{max}} \times b_v)$$

$$Similarity = (0.542 \times 0.25) + (0.333 \times 0.75)$$

$$Similarity = 0.39$$

Sehingga didapatkan nilai kemiripan antara task pertama PMBOK dan Scrum dengan nilai 0.39. Setiap task lainnya dibandingkan antara satu sama lain sehingga menghasilkan sebuah gambar perbandingan sebagai berikut :

	A1							Rata-Rata
		a1	a2	a3	a4	a5	a6	
A2	b1	0.39	0.63	0.53	0.55	0.45	0.55	0.52
	b2	0.43	0.5	0.91	0.62	0.6	0.6	0.61
	b3	0.46	0.49	0.53	0.48	0.37	0.42	0.458
	b4	0.33	0.36	0.53	0.4	0.54	0.51	0.445
	b5	0.9	0.6	0.76	0.37	0.4	0.35	0.563
	b6	0.7	0.57	0.43	0.44	0.6	0.56	0.55
								0.52

Gambar 8. Hasil perhitungan rata-rata tiap task

Setiap task dalam knowledge dibandingkan sehingga menghasilkan matrik besar pada gambar 9 di bawah ini :

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10
A1	1	0.45	0.51	0.44	0.45	0.45	0.67	0.49	0.45	0.36	0.52	0.47	0.41	0.39	0.5	0.52	0.49	-	-	0.4
A2	0.45	1	0.63	0.39	0.53	0.42	0.55	0.5	0.48	0.48	0.48	0.57	0.51	0.46	0.35	0.3	0.37	-	-	0.38
A3	0.51	0.63	1	0.45	0.55	0.49	0.57	0.53	0.5	0.4	0.3	0.56	0.51	0.5	0.45	0.38	0.3	-	-	0.42
A4	0.44	0.39	0.59	1	0.59	0.74	0.58	0.63	0.47	0.33	0.39	0.28	0.4	0.51	0.3	0.58	0.2	-	-	0.4
A5	0.45	0.53	0.55	0.59	1	0.61	0.69	0.6	0.4	0.43	0.4	0.3	0.43	0.53	0.43	0.53	0.4	-	-	0.36
A6	0.45	0.42	0.49	0.74	0.61	1	0.61	0.66	0.43	0.41	0.35	0.34	0.38	0.48	0.53	0.45	0.3	-	-	0.43
A7	0.67	0.55	0.57	0.74	0.69	0.61	1	0.64	0.39	0.4	0.3	0.4	0.39	0.4	0.35	0.3	0.42	-	-	0.4
A8	0.49	0.5	0.53	0.63	0.6	0.66	0.64	1	0.47	0.43	0.36	0.43	0.46	0.3	0.46	0.4	0.51	-	-	0.35
A9	0.45	0.48	0.5	0.47	0.4	0.43	0.39	0.47	1	0.46	0.39	0.38	0.42	0.5	0.53	0.36	0.59	-	-	0.39
A10	0.36	0.48	0.4	0.33	0.43	0.41	0.4	0.43	0.46	1	0.3	0.3	0.36	0.48	0.39	0.45	0.53	-	-	0.48
B1	0.57	0.48	0.3	0.39	0.4	0.35	0.3	0.36	0.39	0.3	1	0.53	0.61	0.45	0.45	0.41	0.64	-	-	0.62
B2	0.47	0.57	0.56	0.28	0.3	0.34	0.4	0.43	0.38	0.3	0.53	1	0.4	0.38	0.55	0.45	0.53	-	-	0.67
B3	0.41	0.51	0.51	0.4	0.43	0.38	0.39	0.46	0.42	0.36	0.61	0.4	1	0.47	0.6	0.51	0.47	-	-	0.6
B4	0.39	0.46	0.5	0.51	0.53	0.48	0.4	0.3	0.5	0.48	0.45	0.38	0.47	1	0.51	0.49	0.56	-	-	0.49
B5	0.5	0.35	0.45	0.3	0.43	0.53	0.35	0.46	0.53	0.39	0.45	0.55	0.6	0.51	1	0.6	0.51	-	-	0.38
B6	0.52	0.3	0.36	0.58	0.53	0.45	0.3	0.4	0.36	0.45	0.41	0.45	0.51	0.49	0.6	1	0.57	-	-	0.41
B7	0.49	0.37	0.3	0.2	0.4	0.3	0.42	0.51	0.59	0.53	0.64	0.53	0.47	0.56	0.51	0.57	1	-	-	0.47
B8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-
B9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-
B10	0.5	0.34	0.47	0.41	0.38	0.49	0.6	0.67	0.62	0.48	0.39	0.35	0.4	0.43	0.36	0.4	0.42	-	-	1

Gambar 9. Nilai perhitungan nilai *semantic similarity* semua *task*

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 Gambar 10 di bawah ini:

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10
A1	1.00	0.79	0.52	0.53	0.67	0.54	0.72	0.50	0.57	0.65	0.69	0.65	0.42	0.36	0.53	0.58	0.00	0.00	0.44	0.63
A2	0.79	1.00	0.55	0.54	0.71	0.55	0.73	0.51	0.52	0.54	0.68	0.65	0.45	0.42	0.57	0.54	0.00	0.00	0.47	0.66
A3	0.52	0.63	1.00	0.67	0.72	0.66	0.65	0.60	0.62	0.48	0.51	0.58	0.52	0.40	0.68	0.37	0.00	0.00	0.56	0.51
A4	0.53	0.54	0.67	1.00	0.58	0.82	0.57	0.66	0.58	0.42	0.44	0.52	0.60	0.55	0.57	0.42	0.00	0.00	0.64	0.45
A5	0.67	0.71	0.72	0.58	1.00	0.58	0.81	0.47	0.53	0.57	0.56	0.58	0.43	0.46	0.62	0.54	0.00	0.00	0.47	0.63
A6	0.54	0.55	0.75	0.82	0.58	1.00	0.59	0.65	0.56	0.37	0.50	0.49	0.53	0.50	0.44	0.44	0.00	0.00	0.66	0.58
A7	0.72	0.73	0.65	0.57	0.81	0.59	1.00	0.50	0.54	0.58	0.61	0.62	0.33	0.42	0.59	0.66	0.00	0.00	0.43	0.65
A8	0.50	0.51	0.60	0.66	0.47	0.65	0.50	1.00	0.59	0.42	0.49	0.51	0.59	0.60	0.48	0.53	0.00	0.00	0.65	0.45
A9	0.57	0.52	0.62	0.58	0.53	0.56	0.54	0.59	1.00	0.43	0.52	0.56	0.51	0.46	0.61	0.56	0.00	0.00	0.61	0.46
A10	0.65	0.54	0.48	0.42	0.57	0.37	0.58	0.42	0.43	1.00	0.62	0.66	0.40	0.40	0.55	0.79	0.00	0.00	0.55	0.75
B1	0.69	0.68	0.51	0.44	0.56	0.50	0.61	0.49	0.52	0.62	1.00	0.66	0.41	0.50	0.69	0.62	0.00	0.00	0.65	0.58
B2	0.65	0.65	0.58	0.52	0.58	0.49	0.62	0.51	0.56	0.66	0.66	1.00	0.46	0.53	0.72	0.58	0.00	0.00	0.61	0.55
B3	0.42	0.45	0.52	0.60	0.43	0.53	0.33	0.59	0.51	0.40	0.41	0.46	1.00	0.72	0.47	0.46	0.00	0.00	0.58	0.36
B4	0.36	0.42	0.40	0.55	0.46	0.50	0.42	0.60	0.46	0.40	0.50	0.53	0.72	1.00	0.53	0.43	0.00	0.00	0.52	0.42
B5	0.53	0.57	0.68	0.57	0.62	0.44	0.59	0.48	0.61	0.55	0.69	0.72	0.47	0.53	1.00	0.64	0.00	0.00	0.50	0.61
B6	0.58	0.54	0.37	0.42	0.54	0.44	0.66	0.53	0.56	0.79	0.62	0.58	0.46	0.43	0.64	1.00	0.00	0.00	0.46	0.71
B7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-	-	0.00	0.00
B8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-	-	0.00	0.00
B9	0.42	0.50	0.57	0.65	0.51	0.77	0.61	0.78	0.61	0.42	0.47	0.50	0.55	0.51	0.50	0.44	0.00	0.00	1.00	0.48
B10	0.63	0.66	0.51	0.45	0.63	0.58	0.65	0.45	0.46	0.73	0.58	0.55	0.36	0.42	0.61	0.71	0.00	0.00	0.42	1.00

Gambar 10. Hasil perhitungan tiap *task* dengan bobot

Hasil dari pembobotan penghitungan kemiripan pada Gambar 10 kemudian digunakan pada proses *clustering*.

Clustering

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

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10
A1	1.0	0.7	0.5	0.5	0.6	0.5	0.7	0.5	0.5	0.65	0.6	0.6	0.4	0.3	0.5	0.5	0.0	0.0	0.4	0.63
A2	0.7	1.0	0.5	0.5	0.7	0.5	0.7	0.5	0.5	0.54	0.6	0.6	0.4	0.4	0.5	0.5	0.0	0.0	0.4	0.66
A3	0.5	0.6	1.0	0.6	0.7	0.6	0.6	0.6	0.6	0.48	0.5	0.5	0.5	0.4	0.6	0.3	0.0	0.0	0.5	0.51
A4	0.5	0.5	0.6	1.0	0.5	0.8	0.5	0.6	0.5	0.42	0.4	0.5	0.6	0.5	0.5	0.4	0.0	0.0	0.6	0.45
A5	0.6	0.7	0.7	0.5	1.0	0.5	0.8	0.4	0.5	0.57	0.5	0.5	0.4	0.4	0.6	0.5	0.0	0.0	0.4	0.63
A6	0.5	0.5	0.7	0.8	0.5	1.0	0.5	0.6	0.5	0.37	0.5	0.4	0.5	0.5	0.4	0.4	0.0	0.0	0.6	0.58
A7	0.7	0.7	0.6	0.5	0.8	0.5	1.0	0.5	0.5	0.58	0.6	0.6	0.3	0.4	0.5	0.6	0.0	0.0	0.4	0.65
A8	0.5	0.5	0.6	0.6	0.4	0.6	0.5	1.0	0.5	0.42	0.4	0.5	0.5	0.6	0.4	0.5	0.0	0.0	0.6	0.45
A9	0.5	0.5	0.6	0.5	0.5	0.5	0.5	0.5	1.0	0.43	0.5	0.5	0.5	0.4	0.6	0.5	0.0	0.0	0.6	0.46
A10	0.6	0.5	0.4	0.4	0.5	0.3	0.5	0.4	0.4	1.00	0.6	0.6	0.4	0.4	0.5	0.7	0.0	0.0	0.5	0.75
B1	0.6	0.6	0.5	0.4	0.5	0.5	0.6	0.4	0.5	0.62	1.0	0.6	0.4	0.5	0.6	0.6	0.0	0.0	0.6	0.58
B2	0.6	0.6	0.5	0.5	0.5	0.4	0.6	0.5	0.5	0.66	0.6	1.0	0.4	0.5	0.7	0.5	0.0	0.0	0.6	0.55
B3	0.4	0.4	0.5	0.6	0.4	0.5	0.3	0.5	0.5	0.40	0.4	0.4	1.0	0.7	0.4	0.4	0.0	0.0	0.5	0.36
B4	0.3	0.4	0.4	0.5	0.4	0.5	0.4	0.6	0.4	0.40	0.5	0.5	0.7	1.0	0.5	0.4	0.0	0.0	0.5	0.42
B5	0.5	0.5	0.6	0.5	0.6	0.4	0.5	0.4	0.6	0.55	0.6	0.7	0.4	0.5	1.0	0.6	0.0	0.0	0.5	0.61
B6	0.5	0.5	0.3	0.4	0.5	0.4	0.6	0.5	0.5	0.79	0.6	0.5	0.4	0.4	0.6	1.0	0.0	0.0	0.4	0.71
B7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.0	0.0	0.0	0.0	0.0	0.0	-	-	0.0	0.00
B8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.0	0.0	0.0	0.0	0.0	0.0	-	-	0.0	0.00
B9	0.4	0.5	0.5	0.6	0.5	0.7	0.6	0.7	0.6	0.42	0.4	0.5	0.5	0.5	0.5	0.4	0.0	0.0	1.0	0.48
B10	0.6	0.6	0.5	0.4	0.6	0.5	0.6	0.4	0.4	0.73	0.5	0.5	0.3	0.4	0.6	0.7	0.0	0.0	0.4	1.00

Gambar 11. Hasil *clustering* semua task

Pada Gambar 11 menunjukkan bahwa terdapat 2 kluster. Kluster 1 merepresentasikan kelompok proses bisnis dari masing-masing knowledge PMBoK. Sedangkan kluster 2 merepresentasikan kelompok proses bisnis dari masing-masing knowledge Scrum. Menggunakan nilai batas minimal 0.49 dan batas maksimal 0.55, didapatkan sebuah *True Positive* yang berwarna hijau, *True Negative* yang berwarna putih, *False Positive* yang berwarna oranye, dan *False Negative* yang tidak ditemukan pada *clustering* di atas.

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 Setelah menguji semua kombinasi untuk setiap algoritma *similarity*, langkah selanjutnya adalah menghitung akurasi dari setiap hasil dari proses *clustering* yang dijalankan. Tujuan dari proses perhitungan presisi adalah untuk menemukan kombinasi yang tepat, algoritma yang digunakan, dan ambang batas untuk menemukan fragmen umum yang tepat. Metode yang digunakan untuk mengukur akurasi adalah ROC. Hasil ketelitiannya adalah

$$\text{rumus : Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

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

$$Accuracy = \frac{171 + 168}{171 + 168 + 61 + 0} \times 100\%$$

$$Accuracy = 84,75\%$$

KESIMPULAN

Komposisi nilai bobot dari perhitungan *semantic similarity* adalah 0,75 untuk kelas verba dan 0,25 untuk kelas kata benda. Komposisi bobot kesamaan adalah 0,10 untuk perilaku, 0,33 untuk struktur, dan 0,57 untuk semantik. Proses *clustering* dilakukan dengan membandingkan nilai kumulatif hasil perhitungan *similarity weighting* dengan nilai *threshold* yang telah ditentukan. Jika skor kesamaan ($\geq$) adalah ambang batas, skor melewati cluster. Ambang batasnya adalah 0,49. Skor akurasi keseluruhan untuk kesamaan antara kedua model proses bisnis adalah 84,75, dan persamaan tersebut ditunjukkan pada Gambar 11. Dari hasil perhitungan tersebut nantinya dapat digunakan sebagai *benchmark* untuk membuat model proses bisnis.

REFERENSI

- [1] H. F. Cervone, "Understanding agile project management methods using Scrum," *OCLC Syst. Serv. Int. Digit. Libr. Perspect.*, vol. 27, no. 1, pp. 18–22, Jan. 2011, doi: 10.1108/10650751111106528.
- [2] S. Chaouch, A. Mejri, and S. A. Ghannouchi, "A framework for risk management in Scrum development process," *Procedia Comput. Sci.*, vol. 164, pp. 187–192, Jan. 2019, doi: 10.1016/J.PROCS.2019.12.171.
- [3] J. Magano, G. Silvius, C. S. e Silva, and A. Leite, "The contribution of project management to a more sustainable society: Exploring the perception of project managers," *Proj. Leadersh. Soc.*, vol. 2, p. 100020, Dec. 2021, doi: 10.1016/j.plas.2021.100020.
- [4] S. W. van Rooij, "Scaffolding project-based learning with the project management body of knowledge (PMBOK®)," *Comput. Educ.*, vol. 52, no. 1, pp. 210–219, Jan. 2009, doi: 10.1016/j.compedu.2008.07.012.
- [5] S. K. Sari and A. Asniar, "Analisis Dan Pemodelan Proses Bisnis Prosedur Pelaksanaan Proyek Akhir Sebagai Alat Bantu Identifikasi Kebutuhan Sistem," *J. Infotel*, vol. 7, no. 2, pp. 143–152, 2015.
- [6] D. Faturrachman, D. S. Kusumo, and K. R. S. Wiharja, "Analisis Aspek *Behavioral* Pada Business Process Model And Notation Menggunakan Causal Footprints," *EProceedings Eng.*, vol. 2, no. 3, 2015.
- [7] G. U. Abriani, M. A. Yaqin, and F. Fatchurrohman, "Pengukuran Kemiripan Model Proses Bisnis untuk menentukan Common Fragment," *MATICS*, vol. 12, no. 2, pp. 117–123, 2020.
- [8] M. K. Shenoy, K. Shet, and U. D. Acharya, "A new *similarity* measure for taxonomy based on edge counting," *Int. J. Web Semantic Technol.*, vol. 3, no. 4, p. 23, 2012.
- [9] M. N. Arifin and D. Siahaan, "Structural and Semantic Similarity Measurement of UML Use Case Diagram," *Lontar Komput. J. Ilm. Teknol. Inf.*, vol. 11, no. 2, p. 88, 2020.
- [10] M. A. Tayal, M. M. Raghuwanshi, and L. Malik, "Word net based method for determining *semantic sentence similarity* through various word senses," in *Proceedings of the 11th international conference on natural language processing*, 2014, pp. 139–145.
- [11] Y. Zhou and H. Cheng, "Je rey Xu Yu. 2009," *Graph Clust. Based Struct. Ribute Similarities VLDB*, 2009.
- [12] M. Sokolova, N. Japkowicz, and S. Szpakowicz, "Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation," in *AI 2006: Advances in Artificial Intelligence*, Berlin, Heidelberg, 2006, pp. 1015–1021.
- [13] T. Fawcett, "An introduction to ROC analysis," *ROC Anal. Pattern Recognit.*, vol. 27, no. 8, pp. 861–874, Jun. 2006, doi: 10.1016/j.patrec.2005.10.010.