

Analisis Sentimen Opini Masyarakat Terhadap Pasangan Calon Presiden dan Wakil Presiden pada Media Sosial Twitter Menggunakan *Ontology Supported Polarity Mining* (Studi Kasus: Pemilihan Umum 2019)

Yustisia Susandi¹, Anisa Herdiani², Indra Lukmana Sardi³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹yustisiachan@students.telkomuniversity.ac.id, ²anisaherdiani@telkomuniversity.ac.id,

³indraluk@telkomuniversity.ac.id

Abstrak

Pemilihan umum (pemilu) presiden dan wakil presiden tak lepas dari pembicaraan publik. Masyarakat bebas mengemukakan opininya baik melalui media sosial maupun di dunia nyata. Salah satu media sosial yang digemari masyarakat untuk menyampaikan opini adalah Twitter. Banyaknya masyarakat yang beropini mengenai pemilihan presiden di Twitter dapat menghasilkan informasi yang dapat dimanfaatkan lembaga survei untuk menentukan tingkat keterpilihan atau elektabilitas pasangan calon presiden dan wakil presiden. Oleh karena itu, perlu adanya analisis sentimen opini masyarakat tentang Pemilihan Umum 2019 yang berkaitan dengan para pasangan calon, sehingga dapat diketahui bagaimana penilaian masyarakat terhadap pasangan calon presiden dan wakil presiden yang sudah diklasifikasikan ke dalam masing-masing aspek secara spesifik. Metode yang digunakan untuk analisis sentimen opini masyarakat adalah *Ontology Supported Polarity Mining* (OSPM). Metode OSPM dipilih karena dapat meningkatkan penambangan polaritas dengan ontologi yang dapat memberikan informasi suatu topik secara terperinci. Hasil menunjukkan bahwa OSPM dengan teknik penambangan polaritas *rule* dapat mengklasifikasikan sentimen level aspek dan menghasilkan akurasi sebesar 93.76% sedangkan tanpa *rule* rata-rata akurasi sebesar 85.11%, dengan kata lain *rule* dapat meningkatkan akurasi sebesar 8.65%.

Kata Kunci : sentimen, level aspek, OSPM, pilpres, *twitter*

Abstract

The general election (election) of the president and vice president cannot be separated from public discussion. People are free to express their opinions both through social media and in the real world. One of the social media favored by the public to express their opinion is Twitter. The number of people who think about the presidential election on Twitter can produce information that can be used by survey institutions to determine the level of electability or electability of candidates for president and vice presidential candidates. Therefore, it is necessary to analyze the sentiment of public opinion about the 2019 General Election relating to the candidate pairs, so that it can be seen how the community evaluates the presidential and vice presidential candidate pairs that have been specifically classified into each aspect. The method used for public opinion sentiment analysis is *Ontology Supported Polarity Mining* (OSPM). The OSPM method was chosen because it can improve polarity mining with ontologies that can provide detailed information on a topic. The results show that OSPM with polarity rule mining techniques can classify sentiment level aspects and produce an accuracy of 93.76% while without rule the average accuracy is 85.11%, in other words the rule can increase accuracy by 8.65%.

Keywords : sentiment, aspect level, OSPM, presidential election, twitter

1. Pendahuluan

1.1 Latar Belakang

Pemilihan umum (pemilu) merupakan sarana bagi rakyat untuk ikut serta menentukan figur dan arah kepemimpinan negara dalam periode waktu tertentu [1]. Pemilu untuk pemilihan presiden (pilpres) dan wakil presiden di Indonesia diselenggarakan setiap lima tahun sekali. Pada tahun 2019 pemilihan presiden dan wakil presiden diselenggarakan yang keempat kalinya [2]. Pembicaraan mengenai pemilu tersebut tidak dapat dipungkiri. Masyarakat bebas mengemukakan opininya baik melalui media sosial maupun di dunia nyata [3]. Menurut Peneliti Pusat Penelitian Politik, Lembaga Ilmu Pengetahuan Indonesia (LIPI) Wasisto Raharjo Jati mengatakan, bahwa media sosial sering digunakan untuk mendominasi opini masyarakat mengenai pemilihan umum, salah satunya adalah Twitter [4]. *Tweet* tentang pilpres yang berjumlah ribuan data per harinya memiliki batasan, yaitu hanya 280 karakter per-*tweet* [5]. Dengan demikian, opini masyarakat yang dikemukakan di

Twitter pun biasanya ditulis tidak mengikuti standar baku tata bahasa, terdapat singkatan-singkatan dan simbol-simbol tertentu yang digunakan sebagai wujud pengekspresian pendapatnya [6]. Hal ini menimbulkan tantangan penelitian apabila dilakukan secara manual untuk mengolah opini masyarakat tersebut agar menjadi sebuah informasi karena dibutuhkan waktu dan usaha yang lebih banyak [6].

Adapun sebuah bidang riset, yaitu analisis sentimen yang difokuskan pada studi komputasi atas pendapat, sikap, dan emosi yang dituliskan dalam bentuk teks [7]. Sasarannya adalah untuk membedah opini dan memperoleh polaritas atau klasifikasi atas sentimen yang ada dalam opini tersebut. Secara umum analisis sentimen dibagi menjadi tiga level, yaitu level dokumen, level kalimat, dan level aspek. Pada penelitian ini menggunakan analisis sentimen level aspek. Analisis sentimen level aspek menunjukkan performa yang lebih baik dibandingkan level dokumen dan level kalimat. Hal ini dikarenakan dalam suatu opini yang dikemukakan seseorang cenderung membahas tiap aspek suatu entitas, bukan keseluruhan [8]. Oleh karena itu, perlu adanya analisis sentimen opini masyarakat level aspek agar dapat diketahui bagaimana penilaian masyarakat terhadap pasangan calon presiden dan wakil presiden.

Metode yang digunakan untuk analisis sentimen opini masyarakat ini adalah *Ontology Supported Polarity Mining* (OSPM). Metode OSPM dipilih karena dapat meningkatkan penambangan polaritas dengan ontologi yang dapat memberikan informasi suatu topik secara terperinci [9]. Fungsi dari penambangan polaritas adalah untuk menentukan orientasi positif dan negatif pada sebuah informasi berdasarkan teknik pembelajaran mesin [10]. Pada penambangan polaritas terdapat proses penambangan opini yang bertujuan untuk mengekstrak dan membedah opini yang dikemukakan oleh masyarakat, sehingga dapat meningkatkan pemahaman dan maksud dari opini [10]. Selain itu, metode ini dipilih karena proses klasifikasi dan analisis sentimen menjadi lebih efisien dan sederhana karena data dipisahkan berdasarkan domain tertentu yang telah didefinisikan sebelumnya [9].

1.2 Topik dan Batasannya

Berdasarkan latar belakang yang sudah dijelaskan sebelumnya topik yang dibahas pada penelitian ini adalah sebagai berikut.

- Bagaimana klasifikasi sentimen opini masyarakat level aspek terhadap setiap pasangan calon presiden dan wakil presiden pada Pemilihan Presiden 2019 dengan menggunakan *Ontology Supported Polarity Mining* (OSPM)?
- Bagaimana performansi klasifikasi sentimen level aspek dengan menggunakan *Ontology Supported Polarity Mining* (OSPM)?

Adapun batasan masalah pada penelitian ini adalah sebagai berikut.

- Dataset* yang digunakan berasal dari *tweet* media sosial Twitter.
- Data yang digunakan hanya *tweet* yang berbahasa Indoensia.
- Pengambilan data berdasarkan beberapa *keyword*, yaitu: jokowi, prabowo, kiyai_marufamin, sandiuno.
- Aspek yang digunakan untuk mengklasifikasikan *tweet* di sosial media Twitter adalah kompetensi, integritas, empati, dan representasi sosial [11].
- Jenis *dataset* yang diambil dari *tweet* hanya berupa teks, untuk *web link*, gambar, dan video tidak diambil.
- Ada dua *dataset* digunakan, yaitu *dataset* pasangan calon Jokowi-Ma'ruf dan *dataset* pasangan calon Prabowo-Sandiaga.

1.3 Tujuan

Pada penelitian ini diimplementasikan pengklasifikasian sentimen opini masyarakat level aspek terhadap setiap pasangan calon presiden dan wakil presiden menggunakan *Ontology Supported Polarity Mining* (OSPM). Data diambil dari *tweet* yang berhubungan dengan pembicaraan masyarakat mengenai pasangan calon presiden dan wakil presiden. Dari data *tweet* tersebut diklasifikasikan ke dalam 4 aspek berserta polaritas sentimennya.

1.4 Sistematika Penulisan

Urutan penulisan laporan ini adalah sebagai berikut : Bagian 2 menunjukkan penelitian-penelitian terkait dengan tugas akhir ini. Sistem yang akan diajukan Analisis Sentimen Opini Masyarakat Terhadap Pasangan Calon Presiden dan Wakil Presiden pada Media Sosial Twitter Menggunakan *Ontology Supported Polarity Mining* (Studi Kasus: Pemilihan Umum 2019) akan dijelaskan pada bagian 3. Pada bagian 4 akan didiskusikan mengenai hasil pengujian dan evaluasi sistem. Dan kesimpulan akan dijelaskan pada bagian 5.

2. Studi Terkait

2.1 Kajian Penelitian Sebelumnya

Kajian penelitian tentang opini masyarakat mengenai Pemilihan Presiden 2019 yang terkait dengan pasangan calon pada dasarnya menganalisis sekumpulan data dalam mengklasifikasikan opini ke dalam aspek

dan sentimen tertentu. Ada beberapa penelitian yang telah dilakukan terkait dengan pasangan calon dengan menggunakan data Twitter. Penelitian tersebut dilihat pada Tabel 2.

Tabel 2. Penelitian Sebelumnya

Peneliti	Judul	Metode	Data	Hasil
Faishol Nurhuda, Sari Widya Sihwi, Afrizal Doewes (UNS, 2013)	Analisis Sentimen Masyarakat terhadap Calon Presiden Indonesia 2014 Berdasarkan Opini dari Twitter [12]	Naïve Bayes Classifier	<i>Public timeline tweet</i> Bahasa Indonesia yang merupakan hasil pencarian berdasarkan percakapan seseorang terhadap akun resmi pasangan calon.	Pada penelitian ini klasifikasi sentimen dilakukan hanya pada level kalimat. Hasil pada penelitian ini adalah pasangan Prabowo-Hatta Rajasa mendapatkan 47,7% sentimen positif, 26,4% sentimen negatif, dan 25,9% sentimen netral. Sedangkan untuk pasangan Jokowi-Jusuf Kalla mendapatkan 37,6% sentimen positif, 34,4% sentimen negatif, dan 27,9% sentimen netral.
Ghulam Asrofi Buntoro (Universitas Muhammadiyah Ponorogo, 2016)	Analisis Sentimen Calon Gubernur DKI Jakarta 2017 di Twitter [13]	- Naïve Bayes Classifier - Support Vector Machine	Data <i>tweet</i> yang diambil adalah hanya data yang berbahasa Indonesia berdasarkan kata kunci yang sudah ditentukan.	Pada penelitian ini klasifikasi sentimen dilakukan hanya pada level kalimat. Penelitian ini juga hanya berfokus mengenai akurasi dari masing-masing metode yang digunakan. Klasifikasi sentimen yang menggunakan metode Naïve Bayes memiliki akurasi lebih tinggi dibandingkan dengan Support Vector Machine.

2.2 Ontologi

Ontologi merupakan suatu cara untuk memaparkan pengetahuan atas sebuah konsep dan hubungannya dengan konsep lain sehingga tercipta keterkaitan hubungan yang bisa direpresentasikan dalam bentuk pohon pengetahuan [14]. Secara khusus, sebuah ontologi menyebutkan konsep domain dan hubungan antar konsep serta memberikan dasar semantik yang baik dari deskripsi *machine-understandable* konten digital [10]. Ontologi berguna untuk *polarity mining* karena dapat mengidentifikasi ulasan teks secara terperinci beserta dengan maknanya [10]. Ontologi memiliki beberapa komponen utama dalam pembentukan ontologi, yaitu sebagai berikut [10].

- Class*, merupakan komponen utama dalam pembentukan ontologi. *Class* dapat merepresentasikan kumpulan dari beberapa individu atau *instance* pada sebuah konsep. Contoh kelas misalnya “Hewan” yang merupakan sebuah representasi dari individu-individu seperti “kelinci, kucing, singa”.
- Instance*, merupakan sebuah hal yang paling dasar dalam pembentukan ontologi karena *instance* dapat menjelaskan ruang lingkup yang dibahas dalam sebuah ontologi.
- Relation*, merupakan hubungan yang menjelaskan keterkaitan antar *instance* atau keterkaitan antar *class*.

Sedangkan untuk langkah-langkah pembangunan ontologi adalah sebagai berikut [14].

- Penentuan domain dan *scope* dalam ontologi.
- Pertimbangan ontologi yang sudah ada.
- Penulisan istilah-istilah penting yang berhubungan dengan domain pada ontologi.
- Pendefinisian kelas dan hierarki yang akan digunakan.
- Pendefinisian properti dari kelas yang telah didefinisikan.
- Pendefinisian *facet* dari properti.
- Pembuatan *instance*.

2.3 Polarity Mining

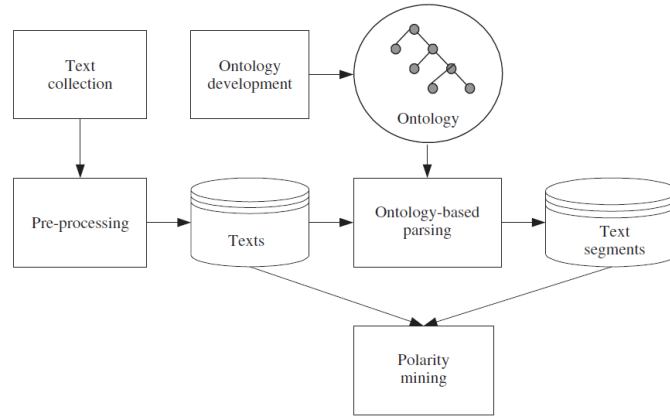
Polarity Mining atau penambangan polaritas adalah suatu cara untuk mengekstrak dan membedah opini yang dikemukakan pada komentar atau penilaian terhadap sesuatu. Dalam penambangan polaritas, fitur merupakan salah satu langkah terpenting untuk mendeteksi kebenaran sentimen contohnya adalah *bag-of-feature*. Beberapa tahapan yang dilakukan untuk memperoleh hasil penambangan polaritas adalah sebagai berikut [10].

- Mengekstrak kata atau frasa yang mengungkapkan orientasi semantik pada sebuah teks.
- Menentukan polaritas kata atau frasa yang diekstrak.
- Menghitung polaritas seluruh kata atau frasa.

Teknik penambangan polaritas yang digunakan adalah *rule-based classifier*, sebuah teknik yang terdiri dari *antecedent* dan *consequent* yang terkait dengan hubungan “if-else”. Antecedent mendefinisikan suatu kondisi dan terdiri dari token atau urutan token yang disambung oleh operator. Token dapat berupa kata ‘?’ mewakili kata benda yang tepat, atau ‘#’ mewakili istilah target. Consequent mewakili suatu sentimen yang positif atau negatif, dan merupakan hasil dari memenuhi kondisi yang didefinisikan oleh antecedent.

2.4 *Ontology Supported Polarity Mining (OSPM)*

Polarity Mining adalah tugas *hybrid* yang ditempatkan antara kategorisasi topik dan sentimen. Manusia dapat dengan sangat mudah mendeteksi kebenaran sentimen pada suatu kata, namun *bag-of-feature classifier* akan sulit untuk menemukan tugasnya karena ada banyak kata yang terindikasi dari sentimen pada seluruh teks dan belum tentu menjadi bagian dari *bag-of-feature classifier*. Penggunaan ontologi memiliki potensi untuk dapat memperbaiki dan meningkatkan proses *Polarity Mining* dengan mengidentifikasi sifat khusus dari sebuah domain [10]. Pada penelitian sebelumnya, pendekatan OSPM digunakan untuk melakukan penilaian terhadap *brand review* dan *movie review*. Arsitektur dari *Ontology Supported Polarity Mining* (OSPM) dapat dilihat pada Gambar 1:



Gambar 1 Arsitektur *Ontology Supported Polarity Mining* [10]

2.5 *Evaluasi Performansi*

Perhitungan presisi dan akurasi ada empat nilai yang digunakan yaitu TP (*True Positive*) jika prediksi positif dan nilai sebenarnya juga positif, TN (*True Negative*) jika data jika prediksi negatif dan nilai sebenarnya juga negatif, FP (*False Positive*) jika prediksi positif dan nilai sebenarnya negatif, dan FN (*False Negative*) jika prediksi negatif dan nilai sebenarnya positif.

Empat nilai di atas akan digunakan untuk perhitungan *accuracy*, *precision* dan *recall*.

Precision adalah tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh sistem dapat dilihat pada persamaan di bawah ini [15].

$$Precision = \frac{TP}{TP + FP} \quad (2.1)$$

Recall adalah perbandingan jumlah data yang diklasifikasikan true secara tepat terhadap seluruh jumlah kelas yang dianggap *true* [16]. Perhitungan *recall* dapat dilihat pada persamaan di bawah ini.

$$Recall = \frac{TP}{TP + FN} \quad (2.2)$$

F1-Measure merupakan hasil rata-rata antara *precision* dan *recall*. *F-measure* merupakan pengukuran performansi sistem yang menggabungkan nilai *precision* dan *recall* [16]. Perhitungan *F1-Measure* dapat dilihat pada persamaan di bawah ini.

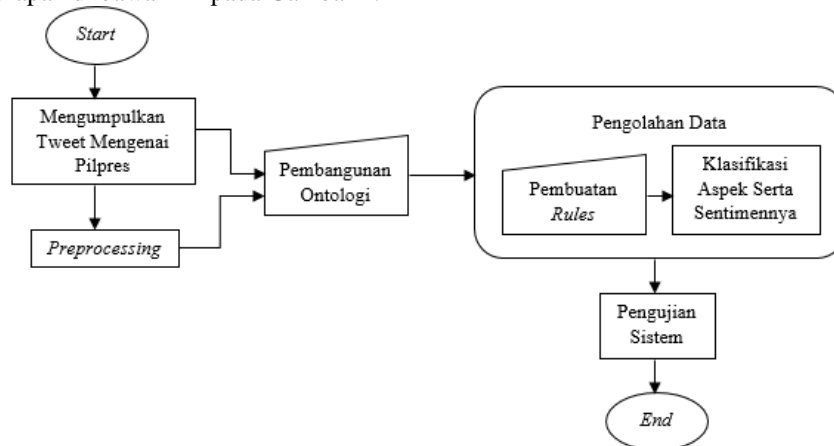
$$F1 - Measure = \frac{2 * precision * recall}{precision + recall} \quad (2.3)$$

Accuracy didefinisikan sebagai tingkat kedekatan antara nilai prediksi dengan nilai aktual. Perhitungan *Accuracy* dapat dilihat pada persamaan di bawah ini.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (2.4)$$

3. Sistem yang Dibangun

Metodologi penelitian yang digunakan untuk membangun sistem dalam penelitian ini digambarkan dengan bagan tahapan di bawah ini pada Gambar 2.



Gambar 2 Alur Pembangunan Sistem

Berikut adalah penjelasan mengenai tahapan-tahapan metodologi yang digunakan.

3.1 Pengumpulan Data

Pada tahap ini *dataset* diambil dari opini publik di Twitter yang membicarakan tentang Pemilihan Presiden 2019 dan berkaitan dengan para pasangan calon. Pasangan calon pada pemilu 2019 itu sendiri ada dua, yaitu Jokowi-Ma'ruf dan Prabowo-Sandi. Proses pengambilan data pada Twitter atau *crawling data* ini memiliki beberapa tahap, yaitu pembuatan akun *Dev Twitter* dan *Apps Twitter* untuk memperoleh akses token dan *consumer key*, serta pembuatan kode program untuk *crawling data* menggunakan Bahasa Pemrograman Python. Data yang diambil hanya yang menggunakan Bahasa Indonesia. Pada penelitian ini digunakan dua *dataset*, yaitu *dataset* untuk pasangan calon Jokowi-Ma'ruf sebanyak 300 dan *dataset* pasangan calon Prabowo-Sandiaga sebanyak 300.

3.2 Preprocessing

Pada tahapan ini dilakukan proses *preprocessing* terhadap data *tweet* mengenai pemilihan umum. *Preprocessing* adalah proses pengubahan bentuk data yang terstruktur sembarang menjadi data yang terstruktur sesuai dengan kebutuhan untuk proses dalam *text mining* [12]. Tahap-tahap pada proses *preprocessing* adalah sebagai berikut [12].

a. Case Folding

Pada tahap ini semua huruf dalam dokumen diubah menjadi huruf kecil.

Tabel 4. Contoh *Case Folding*

Sebelum	Sesudah
Mari kita berkampanye dari keluarga dan orang - orang terdekat kita ajak mereka memilih pemimpin baru karna nyatanya pertahanan sudah nggak sanggup memimpin negara ini ,kegaduhan,persekusi,kenaikan harga ,penistaan agama, ekonomi gagal, hoax,maraknya TKW,	mari kita berkampanye dari keluarga dan orang - orang terdekat kita ajak mereka memilih pemimpin baru karna nyatanya pertahanan sudah nggak sanggup memimpin negara ini ,kegaduhan,persekusi,kenaikan harga ,penistaan agama, ekonomi gagal, hoax,maraknya tkw,

b. Symbol and Number Removal

Penghapusan semua simbol dan nomor dilakukan pada tahap ini.

Tabel 5. Contoh *Symbol and Number Removal*

Sebelum	Sesudah
mari kita berkampanye dari keluarga dan orang - orang terdekat kita ajak mereka memilih pemimpin baru karna nyatanya pertahanan sudah nggak sanggup memimpin negara ini ,kegaduhan,persekusi,kenaikan harga ,penistaan agama, ekonomi gagal, hoax,	mari kita berkampanye dari keluarga dan orang orang terdekat kita ajak mereka memilih pemimpin baru karna nyatanya pertahanan sudah nggak sanggup memimpin negara ini kegaduhan persekusi kenaikan harga penistaan agama ekonomi gagal hoax

hoax,maraknya tkw,	maraknya tkw
--------------------	--------------

c. *Stopword Removal*

Stopword removal adalah tahap yang dilakukan proses penghapusan kata yang tidak bermakna. Pada penelitian ini akan dibuat daftar kata yang sering muncul dan tidak bermakna.

Tabel 6. Contoh *Stopword Removal*

Sebelum	Sesudah
mari kita berkampanye dari keluarga dan orang orang terdekat kita ajak mereka memilih pemimpin baru karna nyatanya pertahanan sudah nggak sanggup memimpin negara ini kegaduhan persekusi kenaikan harga penistaan agama ekonomi gagal hoax maraknya tkw	sudah nggak sanggup memimpin negara ini kegaduhan persekusi kenaikan harga penistaan agama ekonomi gagal hoax maraknya tkw

d. *Non-Standard Word Handling*

Pada tahap ini berfungsi untuk mengubah kata atau frasa informal seperti singkatan diubah menjadi kata atau frasa formal. Pada penelitian ini dibuat kamus kata untuk mengubah kata informal menjadi formal. Pada penelitian ini akan dibuat daftar kata yang akan diubah.

Tabel 7. Contoh *Non-Standard Word Handling*

Sebelum	Sesudah
sudah nggak sanggup memimpin negara ini kegaduhan persekusi kenaikan harga penistaan agama ekonomi gagal hoax maraknya tkw	tidak sanggup memimpin negara kegaduhan persekusi kenaikan harga penistaan agama ekonomi gagal hoax maraknya tkw

e. *Tokenizing*

Pemecahan atau pemisahan kalimat berdasarkan tiap kata yang menyusunnya dilakukan pada tahap ini.

Tabel 8. Contoh *Tokenizing*

Sebelum	Sesudah
tidak sanggup memimpin negara kegaduhan persekusi kenaikan harga penistaan agama ekonomi gagal hoax maraknya tkw	tidak, sanggup, memimpin, negara, kegaduhan, persekusi, kenaikan, harga, penistaan, agama, ekonomi, gagal, hoax, maraknya, tkw

f. *Phrases Lookup*

Phrases lookup adalah tahap melihat apakah ada frasa pada hasil token. Jika ada maka token-token tersebut akan digabungkan menjadi satu.

Tabel 9. Contoh *Phrases Lookup*

Sebelum	Sesudah
tidak, sanggup, memimpin, negara, kegaduhan, persekusi, kenaikan, harga, penistaan, agama, ekonomi, gagal, hoax, maraknya, tkw	tidak_sanggup_memimpin_negara, kegaduhan_persekusi, kenaikan_harga, penistaan_agama, ekonomi_gagal, hoax, maraknya_tkw

3.3 Pembangunan Ontologi

Berikut adalah tahapan pembangunan ontologi yang mengacu pada penelitian “*Ontology Development 101: Guide to Creating Your First Ontology*” oleh Natalya F. Noy dan Deborah L McGuinness [14].

a. Penentuan domain dan *scope* dalam ontologi.

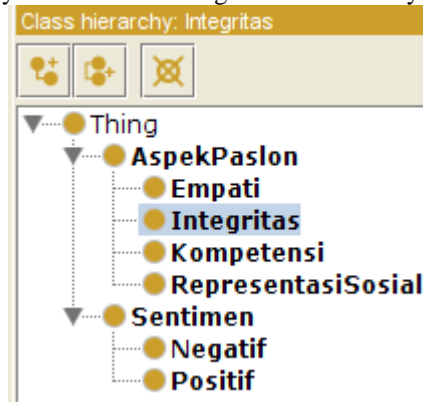
Domain untuk ontologi yang dibangun adalah Aspek Pasangan Calon (Paslon) dan Sentimen.

b. Pertimbangan ontologi yang sudah ada.

Berdasarkan hasil pencarian terdapat aspek pasangan calon (paslon) yang diperoleh dari dimensi-dimensi kualitas paslon menurut hasil survei Lembaga Survei Indonesia dan dapat digunakan kembali untuk pembangunan ontologi pada penelitian ini. Aspek pasangan calon (paslon) yang diperoleh adalah sebagai berikut [11].

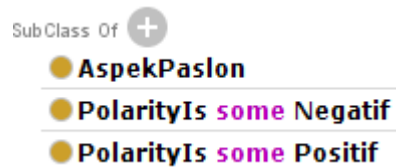
- Kompetensi, dinilai dari kemampuan dalam memimpin dan mengatasi masalah-masalah mendesak yang dihadapi bangsa.
- Integritas, dinilai dari perilaku yang jujur dan dapat dipercaya, kesesuaian antara ucapan dan perbuatan, dan bersih dari korupsi.

- Empati, dinilai dari kepedulian atau perhatian terhadap rakyat.
 - Representasi sosial, dinilai dari agama, asal daerah, kiprah politik, kiprah pendidikan dan karier.
- c. Penulisan istilah-istilah penting yang berhubungan dengan domain pada ontologi.
Istilah-istilah penting yang berhubungan dengan masing-masing domain yang sudah ditentukan untuk domain Pemilihan Presiden 2019 diperoleh empat aspek berdasarkan Lembaga Survei Indonesia, yaitu kompetensi, integritas, empati, dan representasi sosial [11]. Sedangkan, untuk domain sentimen diperoleh dua, yaitu positif dan negatif. Kemudian ditentukan istilah-istilah yang terhubung dengan aspek dan sentimen tersebut yang akan menghasilkan daftar istilah di dalam domain aspek paslon dan sentimen untuk digunakan dalam ontologi.
- d. Pendefinisian kelas dan hierarki yang akan digunakan.
Tahap selanjutnya adalah memetakan istilah-istilah yang sudah terkumpul pada tahap sebelumnya ke dalam kelas-kelas. Terdapat dua kelas yang dibentuk, yaitu AspekPaslon dan Sentimen. Kelas AspekPaslon memiliki empat *sub-class*, yaitu Kompetensi, Integritas, Empati, dan RepresentasiSosial. Sedangkan untuk *sub-class* Sentimen terdapat dua, yaitu Positif dan Negatif. Pendekatan yang digunakan adalah *hybrid*.



Gambar 3 Kelas dan Hierarki

- e. Pendefinisian relasi dan hubungan antar domain.
Mendefinisikan relasi, mempresentasikan relasi atau hubungan antar domain atau konsep. Mendeskripsikan diagram dan relasinya secara detail dengan memberikan nama, domain asal, domain tujuan dan kardinalitas (berapa banyak class dari konsep yang berelasi dengan yang lain), inverse name (inverse dari AspekPaslon ke Sentimen juga sama dengan Sentimen ke AspekPaslon). *Object Properties* untuk menentukan relasi dapat dilihat pada Gambar 4.



Gambar 4 Relasi Aspek Paslon dengan Sentimen

- f. Pembuatan *instance*.
Instance merupakan sebuah objek yang dibuat oleh sebuah *class* [14]. Berdasarkan kelas-kelas yang telah didefinisikan sebelumnya, maka istilah-istilah yang sudah diperoleh dapat dijadikan *instance* dari *sub-class* yang sudah ditentukan. Jumlah *instance* dari masing-masing *subclass* adalah Kompetensi sebesar 140, Integritas sebesar 184, Empati sebesar 71, dan RepresentasiSosial sebesar 120. Berikut adalah salah satu contoh *instance* pada salah satu *subclass*, yaitu Integritas terlihat pada Gambar 5 lebih lengkapnya dapat dilihat pada Lampiran 4.



Gambar 5 Instance pada Sub-Class Integritas

3.4 Pembuatan Rules

Pada penelitian ini penetapan aturan *rule-based* menggunakan aturan *if-then* untuk klasifikasi [17]. Klausa *if* pada aturan *rule-based* merupakan *antecedent* atau *precondition*, dapat terdiri satu atau lebih *token* untuk memperlihatkan kondisi. Untuk klausa *then conclusion* atau *rule consequent* pada aturan *rule-based* merupakan 5 keputusan dari kondisi yang sudah dibuat pada klausa *if*. Untuk penanganan analisis sentimen secara implisit diperlukan rules, termasuk mengatasi dimensional waktu. Aturan yang digunakan pada penelitian ini untuk mengklasifikasikan aspek dan sentimen. Rule yang dibuat berdasarkan hasil validasi dari pakar yang dapat dilihat pada Lampiran 1.

$$R1 = \text{AspekPaslon}(A) \cap \text{Sentimen}(\text{negatif}) \rightarrow \text{PolarityIs}(\text{negatif}) \quad (3.1)$$

$$R2 = \text{AspekPaslon}(A) \cap \text{Sentimen}(\text{positif}) \rightarrow \text{PolarityIs}(\text{positif}) \quad (3.2)$$

$$R3 = \text{AspekPaslon}(\text{'kerja_ambisius'}) \cap \neg \text{Sentimen}(\text{negatif}) \rightarrow \text{PolarityIs}(\text{positif}) \quad (3.3)$$

$$R4 = \text{AspekPaslon}(\text{'kerja_santai'}) \cap \neg \text{Sentimen}(\text{positif}) \rightarrow \text{PolarityIs}(\text{negatif}) \quad (3.4)$$

$$R5 = \text{AspekPaslon}(\text{tidak}(\text{negatif})) \cap \text{Sentimen}(\text{positif}) \rightarrow \text{PolarityIs}(\text{positif}) \quad (3.5)$$

$$R6 = \text{AspekPaslon}(\text{tidak}(\text{positif})) \cap \text{Sentimen}(\text{negatif}) \rightarrow \text{PolarityIs}(\text{negatif}) \quad (3.6)$$

$$R7 = \text{PolarityIs}(\text{positif}) \cap \text{PolarityIs}(\text{negatif}) \rightarrow \text{PolarityIs}(\text{negatif}) \quad (3.7)$$

Keterangan :

R1 : Jika fitur kata atau frasa dalam aspek paslon memiliki sentimen negatif, maka hasil menunjukkan fitur

aspek paslon tersebut memiliki polaritas negatif.

R2 : Jika fitur kata atau frasa dalam suatu aspek paslon memiliki sentimen positif, maka hasil menunjukkan

fitur aspek paslon tersebut memiliki polaritas positif.

R3 : Jika terdapat fitur kata atau frasa 'kerja_ambisius' dalam suatu aspek paslon, maka hasil menunjukkan

polaritas positif, sedangkan untuk kata 'ambisius' selain 'kerja_ambisius' memiliki polaritas negatif.

R4 : Jika terdapat fitur kata atau frasa 'kerja_santai' dalam suatu aspek paslon, maka hasil menunjukkan polaritas negatif, sedangkan untuk kata 'santai' selain 'kerja_santai' memiliki polaritas positif

R5 : Jika fitur kata atau frasa yang memiliki sentimen negatif, apabila sebelumnya terdapat kata 'tidak', maka hasil menunjukkan fitur aspek paslon tersebut memiliki polaritas positif.

R6 : Jika fitur kata atau frasa yang memiliki sentimen positif, apabila sebelumnya terdapat kata 'tidak', maka

hasil menunjukkan fitur aspek paslon tersebut memiliki polaritas negatif.

R7 : Jika fitur kata atau frasa dalam suatu aspek paslon memiliki kedua polaritas positif dan negatif, maka hasil menunjukkan fitur aspek paslon tersebut memiliki polaritas negatif.

* fitur : kumpulan kata atau frasa

3.5 Klasifikasi Aspek dan Sentimennya

Langkah selanjutnya adalah melakukan klasifikasi aspek serta polaritas dari setiap kata atau frasa dengan menggunakan ontologi yang sudah dibangun sebelumnya. Setelah seluruh kata atau frasa sudah mempunyai aspek berserta dengan polaritasnya, langkah selanjutnya adalah menghitung semua jumlah kata atau frasa yang mengandung aspek serta polaritas. Setiap aspek penilaian mempunyai dua kelas sentimen yang dihitung secara terpisah baik positif maupun negatif. Contoh hasil klasifikasi aspek dan sentimen dapat dilihat pada Gambar 6 di bawah ini serta untuk lebih lengkapnya dapat dilihat pada Lampiran 5.

```
periode_bikin_kecewa, .+RepresentasiSosial, Negatif
,orang_baik, .+Integritas, Positif
,goblok, .+Integritas, Negatif
,tiket_mahal, .+Kompetensi, Negatif
,hanya_boneka, .+Kompetensi, Negatif
,belum_terselesaikan, .+RepresentasiSosial, Negatif
,tidak_berkualitas, .+Kompetensi, Negatif
,reformasi_politik, .+RepresentasiSosial, Positif
```

Gambar 6 Contoh Hasil Klasifikasi

Setelah seluruh kata atau frasa sudah mempunyai aspek berserta dengan polaritasnya. Langkah selanjutnya adalah menghitung semua jumlah kata atau frasa yang mengandung aspek serta polaritas terdapat empat aspek berserta polaritasnya berserta jumlah kata atau frasa, untuk persebarannya bisa dilihat pada Tabel 10.

Tabel 10. Jumlah Kata atau Frasa yang Punya Aspek dan Polaritasnya

Pasangan Calon	Kompetensi		Integritas		Empati		Representasi Sosial	
	Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg

Paslon 1	66	46	114	93	39	5	100	35
Paslon 2	56	19	127	114	55	18	49	50
Total	122	65	241	207	94	23	149	85

4. Evaluasi

Setelah sistem selesai dibuat, maka tahap selanjutnya adalah pengujian terhadap sistem. Pengujian ini bertujuan untuk mengetahui seberapa baik sistem yang telah dibuat dengan mengetahui hasil performansi dari sistem. Pengujian metode ini dilakukan pada dataset Paslon1 dan dataset Paslon2 untuk dapat mengklasifikasikan aspek paslon dan sentimennya.

4.1 Hasil Pengujian

Hasil klasifikasi aspek paslon dan sentimen dengan mencocokkan fitur pada Paslon1 dan Paslon 2 dapat dilihat pada Tabel 11.

Tabel 11. Jumlah Kata atau Frasa

Pasangan Calon	Jumlah Fitur Kata/Frasa yang Berhasil Diuji
Paslon 1	498
Paslon 2	488

Berdasarkan jumlah fitur kata atau frasa pada Tabel 11 sebanyak 498 kata/frasa berhasil diuji dari total fitur *preprocessing* pada dataset Paslon 1 dan sebanyak 488 kata/frasa berhasil diuji dari total fitur *preprocessing* pada dataset Paslon 2. Dengan demikian, *Ontology Supported Polarity Mining* berhasil mengklasifikasikan sentimen opini masyarakat level aspek pada 2 paslon sebanyak 986 kata atau frasa. 986 kata atau frasa dapat diklasifikasikan ke dalam empat aspek dan masing-masing aspek memiliki polaritas positif maupun negatif. Jadi kata atau frasa yang berhasil diuji tersebut terbagi menjadi 8 kelas, yaitu Kompetensi Positif, Kompetensi Negatif, Integritas Positif, Integritas Negatif, Empati Positif, Empati Negatif, Representasi Sosial Positif, Representasi Sosial Negatif. Hasil dari klasifikasi sentimen level aspek pada Paslon1 dan Paslon2 bisa dilihat pada Lampiran 5.

Berikutnya adalah hasil performansi uji klasifikasi setiap aspek beserta sentimennya pada 2 paslon dapat dilihat pada Tabel 12 di bawah ini.

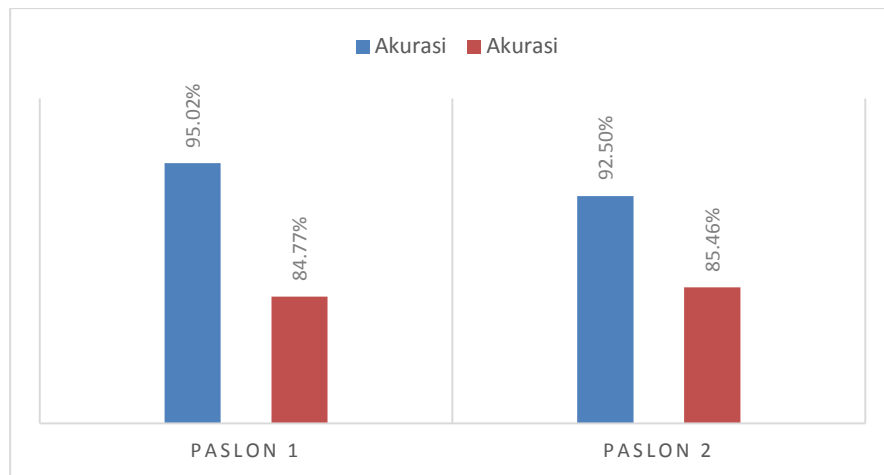
Tabel 12. Hasil Uji Performansi setiap Aspek Paslon beserta Sentimennya

Performansi	Komp Pos	Komp Neg	Int Pos	Int Neg	Emp Pos	Emp Neg	RS Pos	RS Neg
Accuracy	90.17%	95.17%	92.51%	92.51%	94.67%	94.57%	94.34%	96.17%
Precision	85.15%	76.63%	86.94%	85.19%	85.33%	92.17%	75.42%	88.69%
Recall	64.79%	62.38%	84.73%	84.24%	78.30%	48.00%	82.73%	79.75%
F1-Measure	73.52%	68.28%	85.50%	84.63%	81.65%	63.84%	78.38%	83.25%

Berdasarkan Tabel 12, setiap penilaian performansi memiliki presentase yang berbeda-beda pada tiap aspek. Faktor yang menyebabkan menurunnya hasil performansi dijelaskan di sub-bab berikutnya.

4.2 Analisis Hasil Pengujian

Perbandingan hasil akurasi *Ontology Supported Polarity Mining* dengan menggunakan teknik penambangan polaritas *rule-based classifier* dan tanpa menggunakannya dapat dilihat pada Gambar 7.



Gambar 7 Perbandingan Akurasi dengan Rule dan Tanpa Rule

Berdasarkan Gambar 7, apabila hasil akurasi dijumlahkan antara Paslon 1 dan Paslon 2 dengan rule memiliki rata-rata akurasi sebesar 93.76% sedangkan tanpa rule rata-rata akurasi sebesar 85.11%. Dengan demikian, teknik penambangan polaritas dengan menggunakan *rule-based classifier* dapat meningkatkan performansi OSPM sebesar 8.65%. Hal ini menunjukkan bahwa rule dapat menangani beberapa masalah tertentu pada dataset yang memiliki suatu kondisi tertentu. Salah satu contohnya adalah mengatasi kondisional seperti kata ‘ambisius’ umumnya bersentimen negatif, tetapi apabila diikuti dengan kata ‘kerja’ maka ambisius akan bernilai positif. Misalkan ada suatu *tweet*, “jelaslah semua hasil pembangunan nyata, soalnya pakde kan ambisius kerja, ga ambisius nyinyirin orang mulu biar menang.” Kata ‘ambisius’ pada ‘ambisius_kerja’ dalam *tweet* tersebut mengandung polaritas positif, tetapi kata ‘ambisius’ yang sebelumnya diikuti kata ‘ga’ dan ‘nyinyirin’ memiliki polaritas positif.

5. Kesimpulan

Berdasarkan hasil pengujian dan analisis yang telah dilakukan maka dapat diambil kesimpulan sebagai berikut.

- OSPM dapat digunakan untuk mengklasifikasikan aspek serta polaritas opini masyarakat dengan menggunakan teknik penambangan polaritasnya adalah *rule based-classifier*.
- Rata-rata akurasi dengan rule sebesar 93.76% sedangkan tanpa rule rata-rata akurasi sebesar 85.11%. Jadi dapat disimpulkan bahwa penggunaan rule dapat meningkatkan akurasi sebesar 8.65%.

Saran

- Melakukan *preprocessing* yang lebih baik dan lengkap dalam setiap langkah-langkahnya sehingga penanganan data pada tahap selanjutnya dapat lebih baik.
- Membangun ontologi yang lebih lengkap, terutama dari segi *instance class* nya.

Daftar Pustaka

- [1] B. M. Abd Rahman, Pendidikan Kewarganegaraan di Perguruan Tinggi, Makassar: Celebes Media Perkasa, 2017.
- [2] Admin, "Indonesia Memiliki Tradisi Menyelenggarakan Pemilu Secara Periodik," 12 Februari 2016. [Online]. Available: <http://www.kpu.go.id/index.php/post/read/2016/4748/Indonesia-Memiliki-Tradisi-Menyelenggarakan-Pemilu-Secara-Periodik>. [Accessed 13 September 2018].
- [3] M. D. D. Presiden Republik Indonesia, "Undang-Undang Dasar Negara Republik Indonesia Tahun 1945," Jakarta, 2003.
- [4] W. K. Pertiwi, "Kompas Tekno," Kompas, 1 Maret 2018. [Online]. Available: <https://tekno.kompas.com/read/2018/03/01/10340027/riset-ungkap-pola-pemakaian-medsos-orang-indonesia>. [Accessed 13 September 2018].
- [5] Admin, "Kompas Nasional," Kompas, 29 Maret 2014. [Online]. Available: <https://nasional.kompas.com/read/2014/03/29/1153482/Media.Sosial.dalam.Kampanye.Politik>. [Accessed 13 September 2018].
- [6] A. Z. Khan, D. M. Atique and D. V. Thakare, "Combining Lexicon-based and Learning-based Methods for Twitter Sentiment Analysis," *IJECSCSE*, no. 2277-9477, pp. 89-91, 2011.
- [7] B. Liu, Sentiment Analysis and Opinion Mining, Chicago: Morgan & Claypool Publishers, 2012.
- [8] M. Hu and B. Liu, "Mining and Summarizing Customer Reviews," *Journal of Computer Science*, 2012.
- [9] Murugan, B. P.S and A. G, "An Ontology for Exploring Knowledge in Computer Network," *International Journal on Computational Sciences & Applications*, vol. III, 2003.
- [10] P. C. L. Zhou, "Ontology-Supported Polarity Mining," *Journal of The American Society for Information Science and Technology*, vol. 1, no. 98-110, p. 59, 2008.
- [11] L. S. Indonesia, "Isu-isu Paling Mendesak dan Positioning Citra Capres dan Cawapres," PT. Indikator Politik Indonesia, 30 Mei 2012. [Online]. Available: <http://www.lsi.or.id/riset/361/RILIS%20SURVEI%20TERBARU>. [Accessed 31 Desember 2018].
- [12] Fasihol Nurhuda, Sari Widya Sihwi, Afrizal Doewes, "Analisis Sentimen Masyarakat terhadap Calon Presiden Indonesia 2014 Berdasarkan Opini dari Twitter Menggunakan Metode Naive Bayes Classifier," *Jurnal ITSMART*, vol. II, no. 2, 2013.
- [13] G. A. Buntoro, "Analisis Sentimen Calon Gubernur DKI Jakarta 2017 di Twitter," *Integer Journal*, vol. I, no. 1, pp. 32-41, 2016.
- [14] F. N. Natalya and L. M. Deborah, "Ontology Development 101: A Guide to Creating Your First Ontology," Stanford University, CA, 2001.
- [15] G. Cyril and E. Gaussier, "A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation," *Xerox Research Center Europe*, vol. VI, no. F-38240, pp. 345-359, 2005.
- [16] C. D. Manning, P. Raghavan and H. Schutze, Introduction to Information Retrieval, England: Cambridge University Press, 2008.
- [17] R. Prabowo and M. Thelwall, "Sentiment Analysis: A Combined Approach," *Journal of Informatics*, vol. III, no. 2, pp. 143-157, 2009.

