

1. Pendahuluan

1.1 Latar Belakang

Pemilihan umum (pemilu) merupakan salah satu aspek penting dalam sebuah negara demokrasi. Indonesia termasuk dalam negara demokrasi terbesar ketiga di dunia setelah India dan Amerika Serikat [1]. Pemilihan umum di Indonesia dilaksanakan setiap 5 tahun sekali untuk memilih kepala daerah (bupati, gubernur, walikota) maupun pemilihan kepala negara. Gelaran pesta demokrasi beberapa bulan lalu yang hangat diperbincangkan adalah pemilihan gubernur dan wakil gubernur DKI Jakarta 2017 yang diselenggarakan pada tanggal 15 Februari 2017 lalu. Salah satu bagian yang erat kaitannya dengan pemilu adalah survei. Dalam praktek pengambilan survei pemilu, lembaga survei melakukan serangkaian aktivitas survei kepada responden yang meliputi *face to face interview*, wawancara dengan panggilan telepon, survei melalui SMS maupun *online* (internet) [2]. Survei ini menghabiskan waktu dan biaya yang cukup besar sehingga diperlukan alternatif lain yang lebih hemat tanpa harus menanyakan secara langsung kepada masyarakat. Survei ini mengarah pada pemetaan opini publik terhadap kandidat yang mengikuti suatu pemilu. Saat ini penelitian mengenai prediksi kandidat yang memenangi pemilu sudah mulai mengarah pada analisis terhadap opini publik di media sosial, diantaranya Twitter. Analisis pada Twitter dilakukan dengan tujuan untuk mengetahui kecenderungan opini pada *tweet* berdasarkan suatu parameter tertentu. Setiap *tweet* yang dianalisis, akan diklasifikasikan dalam satu jenis kategori opini tertentu. Melalui analisis tersebut, kita mendapatkan data tanpa harus mengeluarkan waktu untuk wawancara langsung dan biaya yang besar sehingga pengambilan data lebih efisien. Untuk itu, pada penelitian ini diusulkan alternatif lain dalam pengumpulan data survei yaitu menggunakan data Twitter dan membuat sistem yang mampu mengenali opini dan memperhitungkan prediksi pemilu.

Berdasarkan data yang dikeluarkan pada Mei 2016 oleh Databoks, Katadata Indonesia [3], pengguna Twitter di Indonesia mencapai 24,34 juta dan menjadikan Indonesia sebagai negara dengan pengguna Twitter terbesar ketiga di dunia setelah Amerika Serikat dan India. Selain itu, pada Juni 2012 SemioCast

menobatkan Jakarta sebagai kota teraktif di dunia dalam *posting tweet*. Jakarta menyumbang lebih dari 2% *tweet* dari 10 miliar *tweet* [4].

Dalam penelitian kali ini, penulis merancang ontologi berdasarkan ragam emosi *tweet* seseorang terhadap pasangan kandidat pilgub DKI Jakarta 2017. Twitter dipilih karena kemudahan pengambilan data dan termasuk salah satu media opini publik sehingga ragam emosi tekstual masyarakat bisa terlihat dari *tweet* yang dibuat. Banyaknya ragam opini publik di Twitter memungkinkan antar opini mengekspresikan makna yang sama. Pendekatan menggunakan ontologi dalam pengklasifikasian *tweet* didasarkan pada kesamaan makna suatu *tweet* yang sehingga lebih menganalisis *tweet* untuk menemukan topik yang terkandung dalam *tweet* tersebut secara terperinci, dalam hal ini berarti mengenai emosi. Parameter emosi yang akan dibangun pada ontologi ini berdasarkan temuan penelitian Shaver and Murdaya [5] yang membagi emosi dalam 5 kelas yaitu *senang, cinta, marah, sedih* dan *takut*. Selain menentukan prediksi hasil pemilu, juga akan dilakukan pengukuran *Mean Absolut Error* (MAE) sebagai evaluasi prediksi untuk melihat perbandingan hasil prediksi dengan hasil yang dikeluarkan oleh masing-masing lembaga survei.

1.2 Perumusan Masalah

Berdasarkan latar belakang tersebut, didapatkan rumusan masalah berupa :

1. Bagaimana skema rancangan ontologi untuk mengklasifikasikan *tweet* dalam menentukan perhitungan prediksi kemenangan pemilu DKI Jakarta 2017?
2. Bagaimana mengklasifikasikan *tweet* berdasarkan parameter emosi dengan menggunakan ontologi?
3. Bagaimana tingkat performansi dari hasil klasifikasi dari ontologi yang dibangun ?
4. Bagaimana mengukur prediksi kemenangan pemilu kandidat DKI Jakarta 2017 berdasarkan hasil klasifikasi ontologi ?

Adapaun batasan masalah yang digunakan dalam penelitian ini adalah :

1. Data *tweet* yang digunakan berada di wilayah provinsi DKI Jakarta pada masa kampanye putaran pertama mulai 12 Desember 2016 hingga 11 Februari 2017.
2. Data *tweet* dari penelitian ini adalah *tweet* berbahasa Indonesia. Data *tweet* yang digunakan berdasarkan *keyword* pencarian.
3. Dalam pembangunan ontologi, ontologi yang digunakan adalah berdasarkan hierarki emosi yang telah dirancang dalam penelitian Shaver and Murdaya dalam “*Structure of Indonesian Emotion Lexicon*”. Pengembangan model ontologi dilakukan dengan mengumpulkan *term* lain terkait parameter emosi yang diambil dari data *tweet*.
4. Penentuan suatu *tweet* ke dalam kelas emosi didasarkan pada kedalaman suatu *term* *tweet* yang cocok pada kelas ontologi. Semakin dalam suatu *term* pada struktur ontologi, semakin spesifik *term* tersebut¹. Dalam kaidah bahasa, semakin sempit cakupan suatu kata, maka semakin spesifik kata tersebut sehingga semakin jelas. *Tweet* akan diklasifikasi pada kelas dengan bobot paling tinggi. Jika didapatkan bobot kelas yang sama, maka klasifikasi didasarkan pada kelas yang pertama kali ditemukan [6]

1.3 Tujuan

Dari rumusan masalah di atas, tujuan dari penelitian ini adalah:

1. Merancang skema ontologi untuk mengklasifikasikan *tweet* dalam menentukan perhitungan prediksi kemenangan pemilu DKI Jakarta 2017.
2. Mengklasifikasikan *tweet* berdasarkan parameter emosi dengan menggunakan ontologi.
3. Menganalisis hasil performansi klasifikasi dari penggunaan ontologi yang telah dibangun.

¹ The lower on the scale, the more concrete and specific the word is –Dikutip dari www.isu.edu/success/writing/handouts/concrete.pdf

4. Mengukur prediksi kemenangan pemilu kandidat DKI Jakarta 2017 berdasarkan hasil klasifikasi ontologi

1.4 Metodologi Penyelesaian Masalah

Metodologi yang digunakan dalam penelitian ini adalah sebagai berikut :

1. Studi literatur

Pada tahapan ini, dilakukan proses pengkajian dan pemahaman materi yang berkaitan dengan penelitian. Kajian materi yang dilakukan adalah mengenai prediksi pemilu dan emosi.

2. Penentuan Parameter

Pada tahap ini dilakukan penentuan parameter emosi yang akan digunakan dalam ontologi. Pada penelitian kali ini, parameter emosi yang digunakan berdasarkan pada parameter emosi yang telah didefinisikan dalam penelitian [5].

3. Pengumpulan Data dan Perancangan Sistem

Pada tahapan ini dilakukan proses pengumpulan data. Data yang digunakan adalah *tweet* yang berada pada wilayah provinsi DKI Jakarta mulai 12 Desember 2016 hingga 11 Februari 2017. Periode tersebut dipilih karena bertepatan dengan masa kampanye. Periode masa tenang kampanye, pengambilan *tweet* tidak dilakukan. *Tweet* yang didapatkan akan dilakukan *text pre-processing*, klasifikasi *tweet* berdasarkan parameter emosi dengan ontologi dan perhitungan prediksi pemilu masing-masing kandidat. Ontologi dibuat secara manual berdasarkan *term-term tweet* dengan aplikasi Protégé.

4. Implementasi Sistem

Sistem dapat melakukan proses pada perancangan sistem seperti praproses, klasifikasi *tweet* berdasarkan ontologi dan perhitungan prediksi.

5. Pengujian dan Analisis

Sistem yang ada akan dilakukan pengujian untuk mendapatkan akurasi terhadap klasifikasi *tweet* berdasarkan ontologi yang dibangun.

6. Kesimpulan dan Pengerjaan Laporan

Seluruh rangkaian proses pengujian yang dilakukan akan ditarik kesimpulan yang sesuai. Semua proses dan tahapan sistem nantinya akan dituang dalam bentuk laporan.

1.5 Sistematika Penulisan

Pada penelitian ini disusun dengan sistematika penulisan sebagai berikut:

1. BAB 1 Pendahuluan

Pada bab ini akan dijelaskan mengenai latar belakang masalah yang diangkat, rumusan dari masalah, tujuan dan metodologi penyelesaian masalah dari penelitian ini.

2. BAB 2 Tinjauan Pustaka

Pada bab ini dijelaskan mengenai teori-teori ataupun kajian penelitian ilmiah lainnya yang berkaitan dengan penelitian yang dilakukan.

3. BAB 3 Perancangan Sistem

Pada bab ini dijelaskan mengenai gambaran umum sistem yang akan dibangun beserta tahapannya.

4. BAB 4 Pengujian dan Analisis

Pada bab ini berisi hasil pengujian yang dilakukan terhadap sistem dan hasil analisis pengujiannya.

5. BAB 5 Kesimpulan dan Saran

Pada bab ini berisi kesimpulan yang didapat dari hasil penelitian dan saran yang dapat digunakan sebagai penyempurnaan penelitian selanjutnya.