

Pendahuluan

I.1 Latar Belakang

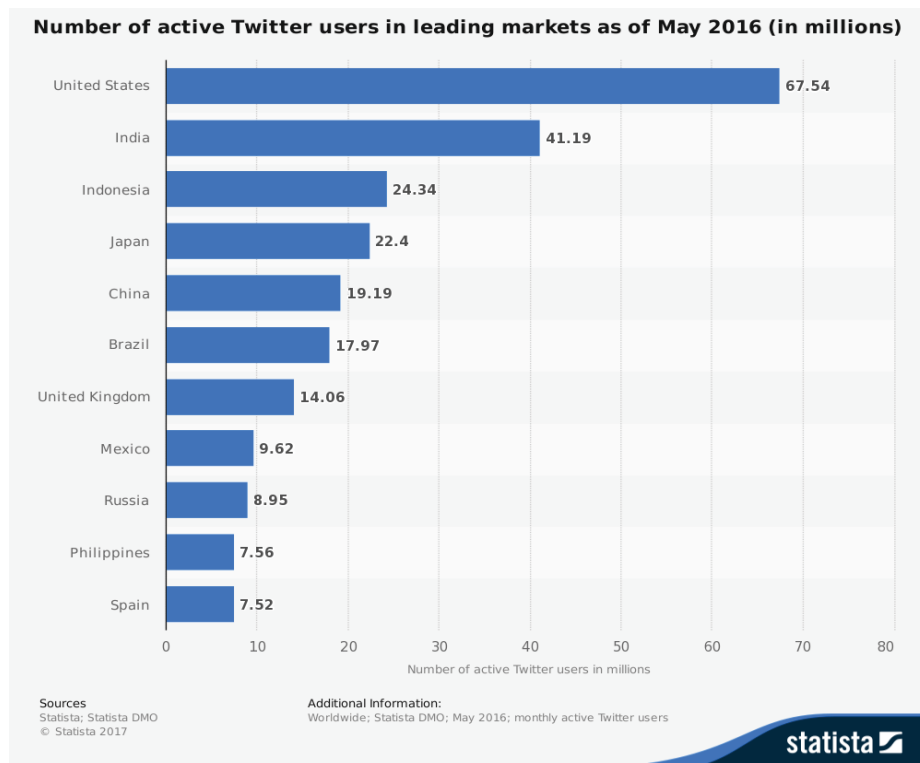
Di Indonesia, pemilihan gubernur adalah acara tiap lima tahun sekali yang diadakan untuk menentukan siapa yang akan memimpin provinsi tersebut selama lima tahun kedepan. Pada tahun 2018 ini, ada 171 daerah yang akan melakukan proses pemilihan kepala daerah yang terdiri dari 17 provinsi, 39 kota, dan 115 kabupaten di Indonesia yang salah satu dari 17 provinsi tersebut adalah provinsi Jawa Barat (Prasetia, 2017). Pada tahun 2018 kali ini di Jawa Barat ada 4 calon kepala daerah yang akan bertarung pada pilkada kali ini, yaitu Ridwan Kamil – UU Ruzhanul Ulum sebagai kandidat nomor urut 1, Tubagus Hasanuddin – Anton Charliyan sebagai kandidat nomor urut 2, Sudrajat – Ahmad Syaikhu sebagai kandidat nomor urut 3, dan Deddy Mizwar – Dedi Mulyadi sebagai kandidat nomor urut 4.

Dengan kemajuan teknologi seperti sekarang ini, maka pasangan calon tersebut harus melakukan kampanye untuk dapat meraih suara rakyat Jawa Barat dalam perjalanan mereka menjadi kepala daerah dari Jawa Barat 2018 – 2023. Berliani Ardha (Ardha, 2014), menunjukkan bahwa partai politik sudah menggunakan media sosial sebagai media kampanye bagi mereka.

Salah satu hal yang kerap dilakukan baik oleh para peneliti maupun pasangan calon kepala daerah pada masa pemilihan adalah melakukan penentuan ataupun prediksi akan elektabilitas dari para pasangan calon kepala daerah tersebut. Tidak jarang Lembaga – lembaga survei ternama memasang angka yang cukup besar dalam pelaksanaan survei tersebut yang membuat baik para peneliti maupun pasangan calon kepala daerah harus mengeluarkan biaya yang sangat besar untuk proses survei elektabilitas tersebut. Menurut Dewan Pimpinan Daerah (DPD) Partai Demokrasi Indonesia Perjuangan (PDIP) Jawa Tengah Bambang Wuryanto pada media Berita Satu menyebutkan bahwa partainya mengeluarkan dana sebesar Rp. 2,6 Miliar untuk melakukan pemilihan kepala daerah (pilkada) di 21 kabupaten/kota (Thenu, 2015). Besarnya biaya survei elektabilitas tersebut, survei dapat menjadi penghalang dalam proses prediksi pemenang proses pemilihan kepala daerah baik

dari pihak peneliti, calon pasangan, dan lainnya. Tingginya biaya akan proses survei inilah yang mendorong adanya kegiatan analisis sentimen yang dilakukan di beberapa media sosial untuk dapat mendapatkan gambaran atau prediksi akan pemenang dari suatu pemilihan kepala daerah. Selain itu, beberapa Lembaga – Lembaga survei seringkali merilis data yang salah terhadap elektabilitas dari calon kepala pemerintahan baik itu presiden, maupun kepala pemerintahan lainnya sehingga hasil analisis sentimen ini dapat menjadi alternatif sumber data yang dapat digunakan untuk mengukur popularitas kandidat calon kepala daerah.

Media sosial merupakan sebuah media berbasis internet yang memungkinkan penggunaanya berkesempatan untuk berinteraksi dan mempresentasikan dirinya baik seketika atau secara tertunda, dengan khalayak luas maupun tidak yang mendorong nilai dari *user-generated content* dan persepsi interaksi dengan orang lain (Carr & Hayes, 2015). Beberapa contoh dari media sosial antara lain seperti *Facebook, Twitter, Instagram, LinkedIn, dan Tumblr*.



Gambar 0.1 Jumlah pengguna Twitter di dunia (Statista, 2017)

Twitter, merupakan salah satu media sosial yang menyediakan layanan *microblogging* bagi para penggunanya. Menurut data yang dirilis dari Statista pada Mei 2016, Indonesia menempati peringkat ketiga dalam jumlah pengguna *Twitter* terbesar di dunia dengan jumlah sebesar 24,3 juta pengguna aktif dibawah India dan Amerika Serikat. Dengan jumlah pengguna yang besar, tak heran jika *Twitter* mendapatkan banyak perhatian dari berbagai kalangan karena kemudahan yang diberikan oleh layanan *microblogging* ini untuk dapat berbagi cerita, pengalaman, dan kejadian yang sedang berlangsung kepada pengguna lainnya dengan cepat. Selain itu, *Twitter* juga kerap menjadi salah satu sarana untuk mengungkapkan perasaan dan emosi dari penggunanya sebagai bagian dari komunikasi manusia, konten – konten dari media sosial seringkali berisikan informasi tentang kondisi emosional, penilaian, evaluasi dari sebuah topik atau seseorang atau sebagai komunikasi emosional, yang sering disebut dengan “sentimen” (Stieglitz & Dang-Xuan, 2013).

Tweets, adalah istilah untuk segala konten yang dibagikan oleh pengguna *Twitter* dalam media sosial tersebut sering kali berisikan informasi – informasi penting yang menjadi fokus penelitian oleh para peneliti terutama di bidang NLP (*Natural Language Processing*). Anshul Mittal dan Arpit Goel melakukan penelitian terhadap hubungan antara data sentimen publik dengan sentimen pasar di *Twitter* dalam menentukan prediksi pergerakan indeks saham pada *Dow Jones Industrial Index* (Mittal & Goel, 2012). Selain itu, Akhsat Bakliwal memberikan gambaran bagaimana melakukan *classifier* yang lebih akurat dalam melakukan analisis sentimen terhadap *tweet* berunsur politis (Bakliwal et al., 2013).

Penggunaan *Twitter* juga semakin menyebar ke segala kalangan baik dari mahasiswa, karyawan, pengusaha, dll. Selain itu dengan banyaknya tokoh – tokoh politik yang memiliki akun di media sosial dengan jumlah pengikut yang sangat besar. Pengikut – pengikut dari tokoh politik inilah yang sering memberikan tanggapan politiknya di *Twitter* baik kepada tokoh yang mereka ikuti, ataupun lawan politik dari tokoh tersebut untuk membantu mengangkat popularitas dan tingkat elektabilitas dari tokoh tersebut.

Twitter, tidak mewajibkan penggunanya untuk menggunakan Bahasa yang baku dalam membagikan *Tweet* sehingga banyak sekali ditemui penggunaan istilah – istilah yang sulit dipahami yang akan menyulitkan proses analisis sentiment. Selain itu, dengan banyaknya data *Tweet* yang ada sistem akan dapat terhambat kinerjanya jika seluruh data *Tweet* baik yang bernilai lebih maupun tidak mengandung nilai diproses secara bersamaan. Oleh karena itu dibutuhkan suatu proses yang dapat mengolah data *Tweet* tersebut siap digunakan dan juga telah dilakukan optimisasi untuk dilakukan proses analisis sentimen secara efektif dan efisien.

Penelitian ini merupakan bagian dari penelitian analisis sentimen bagi kepala daerah yang terbagi ke dalam proses *data retrieval*, *data preprocessing*, dan klasifikasi sentimen.

I.2 Perumusan Masalah

Dari latar belakang yang telah tertera diatas, maka permasalahan yang akan dibahas pada penelitian ini adalah:

1. Pengaruh dari *preprocessing* terhadap jumlah data pada proses analisis sentimen berbasis leksikon.
2. Hasil sentimen dari tiap - tiap pasangan calon kepala daerah Jawa Barat 2018.
3. Pencapaian target performansi dari metode analisis sentimen berbasis leksikon.
4. Korelasi dari hasil analisis sentimen baik dengan hasil survei maupun hasil quick count.

I.3 Tujuan Penelitian

Tujuan dari dilakukannya penelitian ini adalah:

1. Menganalisis pengaruh dari *preprocessing* pada proses analisis sentimen berbasis leksikon terhadap performansi sistem.

2. Melakukan investigasi analisis sentimen dari tiap calon kepala daerah Jawa Barat 2018 yang berupa sentimen positif dan negatif.
3. Menganalisis performa dari metode berbasis leksikon dengan menggunakan *Support Vector Machine*.
4. Membandingkan analisis sentimen yang berdasarkan popularitas dengan hasil survei yang berdasarkan elektabilitas pasangan calon kepala daerah Jawa Barat 2018.

I.4 Manfaat Penelitian

Manfaat yang diberikan dari penelitian ini bagi penulis adalah:

1. Memenuhi syarat untuk mendapatkan gelar sarjana dari Fakultas Rekayasa Industri Telkom University.
2. Memberikan kesempatan untuk dapat menerapkan ilmu yang telah didapat selama studi di Telkom University.

Manfaat yang diberikan dari penelitian ini dalam segi teori adalah:

1. Mengukur performa metode leksikon dengan konteks pemilihan kepala daerah Jawa Barat tahun 2018.
2. Mengetahui korelasi antara elektabilitas dan popularitas dari calon kepala daerah Jawa Barat tahun 2018.

I.5 Ruang Lingkup

Penelitian ini membahas hal – hal yang berkaitan dengan:

1. *Tweet* berbahasa Indonesia.
2. *Tweet* yang berupa teks dengan pengecualian *emoticon*.
3. *Tweet* yang membahas calon gubernur Jawa Barat 2018 dengan mengesampingkan lokasi asal *tweet* tersebut.

I.6 Sistematika Penelitian

Secara garis besar, penyusunan penelitian kali ini dibagi ke dalam beberapa bab sebagai berikut:

1. Bab I Pendahuluan: Berisikan tentang latar belakang masalah, rumusan masalah, tujuan, manfaat serta ruang lingkup atau batasan masalah dari penelitian ini.

2. Bab II Landasan Teori: Berisikan tentang kajian pustaka dari penelitian ini serta penelitian – penelitian yang sudah dilakukan sebelumnya untuk menjadi dasar penelitian ini.
3. Bab III Metodologi Penelitian: Berisikan tentang langkah – langkah dari penelitian ini secara menyeluruh yang mencakup ekstraksi data, *preprocessing* data, serta klasifikasi sentiment.
4. Bab IV Hasil dan Analisis: Berisikan tentang implementasi dari metodologi analisis sentimen, pengukuran performa metode *lexicon based* dengan SVM, serta hasil dari analisis sentimen.
5. Bab V Kesimpulan dan Saran: Berisikan tentang kesimpulan yang didapatkan pada penelitian ini dan saran dari penulis untuk pengembangan penelitian di masa depan.