

BAB I

PENDAHULUAN

1.1. Gambaran Obyek Penelitian

1.1.1. Twitter

Twitter adalah layanan untuk saling terhubung dan berkomunikasi dengan teman-teman, keluarga, dan rekan kerja melalui pertukaran pesan yang cepat dan mudah. Pengguna twitter (*user*) dapat mengunggah pesan (*tweet*) berisi *photo*, *video*, *link* dan teks. Pesan diunggah ke profil *user*, dikirim ke pengikut *user*, dan dapat dicari di pencarian twitter. Pesan twitter bersifat *public* selama *user* tidak mengaktifkan fitur *private account* (Twitter, 2017).



Gambar 1.1. Gambar Logo Twitter

Sumber: Twitter (2017)

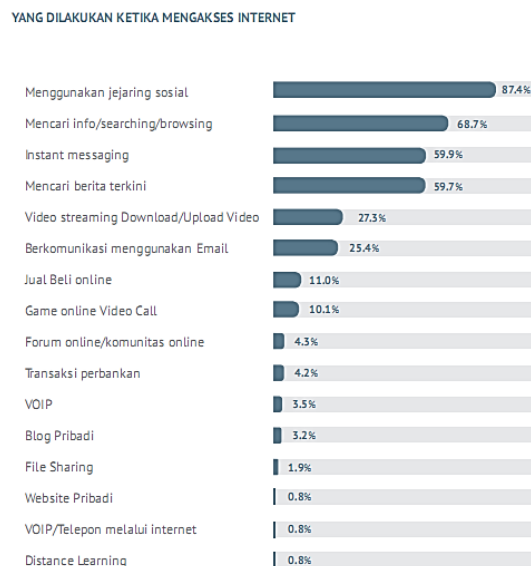
Twitter mempunyai misi untuk memberikan semua orang kekuatan dalam menciptakan dan membagi ide dan informasi secara cepat dan tanpa hambatan. Dengan misi tersebut, berdasarkan data hingga 30 Juni 2016, rata-rata pengguna aktif twitter dapat mencapai 313 Miliar *user* tiap bulan. Sebanyak 79% *user* berasal dari luar Amerika Serikat. Twitter adalah layanan yang dapat digunakan secara global. Saat ini twitter sudah mendukung lebih dari 40 bahasa di seluruh dunia, termasuk bahasa Indonesia (Twitter, 2017).

1.2. Latar Belakang

Media sosial dikenal sebagai kategori wacana online dimana orang membuat konten, membaginya, menyimpannya dan membentuk jaringan yang besar (Asur dan Huberman, 2010). Semua pengguna media sosial mempunyai kesempatan yang sama untuk membuat dan berbagi konten.

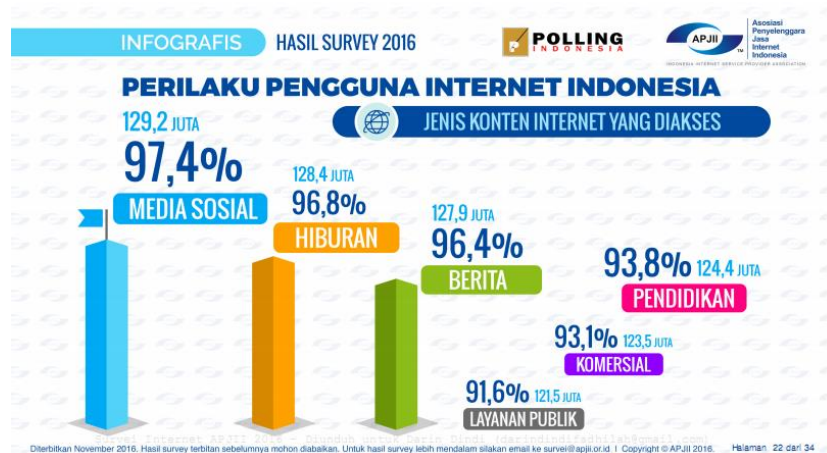
Sebelum adanya media sosial, orang-orang harus bertemu secara personal untuk berinteraksi. Saat ini pengguna telah dimudahkan untuk berinteraksi dengan sesama melalui media sosial. Menurut Java et al (2007), pengguna media sosial mengunggah informasi rutinitas keseharian, melakukan percakapan, melaporkan pemberitaan dan berbagi informasi pada situs media sosialnya. Naaman et al (2010) menambahkan bahwa setiap konten atau interaksi yang diunggah pengguna mengandung keterangan mengenai apa yang mereka rasakan berdasarkan “*social media awareness*”. Pengguna media sosial bebas berekspresi dengan alami, langsung dan terbuka. Hal ini membuat pengguna tidak sadar bahwa teknologi telah merubah akses ke pikiran, perasaan dan aktivitas pengguna (Murphy et al, 2014).

Jumlah pengguna media sosial di Indonesia semakin meningkat. Berdasarkan data survey November 2016 dari Asosiasi Penyelenggara Jaringan Internet Indonesia (APJII), sebanyak 129,9 juta penduduk Indonesia atau 97,4% menggunakan internet untuk media sosial. Jumlah ini meningkat sebesar 10% dari pengguna internet 2014 yang mengakses media sosial.



Gambar 1.2. Yang dilakukan Pengguna Internet Indonesia 2014

Sumber: APJII (2014)



Gambar 1.3. Perilaku Pengguna Internet Indonesia 2016

Sumber: APJII (2016)

Tanpa pengguna sadari, aktivitas pengguna secara langsung akan disimpan sebagai jejak digital pada *database* media sosial. Dewasa ini, para manager melakukan penilaian online terhadap kandidat pegawai dari data jejak digitalnya. Beberapa studi menunjukkan lebih dari 75% manager secara aktif menilai secara online. Dari penilaian tersebut, lebih dari 70% manager berani memutuskan untuk menolak kandidat berdasarkan data yang mereka lihat secara online. Saat ini manager telah mempercayai data dari media sosial untuk mengambil keputusan (Fertik, 2012).

Pada rangkaian seleksi pegawai baru, peserta diharuskan untuk mengikuti tes kepribadian. Alsuwailem dan Elnaga (2016), Sackett et al. (2014), Rothman dan Coetzer (2003) dan Barrick et al. (2001) telah membuktikan adanya pengaruh antara kepribadian dengan performansi kerja pegawai. Performansi kerja pegawai merupakan hal yang penting untuk kesuksesan pegawai di lingkungan kerja. Performansi kinerja dapat menunjukkan seberapa baik karyawan melakukan tugas mereka, inisiatif yang mereka ambil dan akal yang mereka tunjukkan dalam memecahkan masalah.

Tes psikologi kepribadian umumnya dilakukan melalui kegiatan wawancara dan tes tulis. Perusahaan bekerja sama dengan lembaga psikolog

untuk melaksanakan tes dan melakukan penilaian. Pada saat pelaksanaan tes, artikel dari thehrdepartment.ie (2015) mengungkapkan adanya kecenderungan peserta untuk menjawab sesuai dengan keinginan pewawancara atau dengan jawaban yang dapat dinilai baik oleh lingkungan sosial (normatif). Hal tersebut menyebabkan hasil tes tidak dapat mencerminkan kepribadian sesungguhnya.

Barker et al. (1951) dalam Sarwani dan Mahmudy (2015) menyarankan untuk mempelajari kepribadian seseorang pada lingkungan aslinya untuk menghindari perbedaan jawaban dengan memanfaatkan jejak digital. Dengan memanfaatkan kontribusi pengguna media sosial yang sering melakukan *posting* setiap hari, peneliti dapat memanfaatkan jejak digital untuk mengamati aktivitas pengguna dikesehariannya pada lingkungan aslinya. Penelitian lain menyebutkan pula bahwa kepribadian dapat dilihat melalui percakapan sehari-hari (Mehl et al, 2006), serta cara orang menulis dan berbicara dilihat dari penggunaan kata-kata linguistiknya (Stemmler, 2010).

Saat ini banyak penelitian yang menggunakan data dari media sosial sebagai sumber data penelitian. Penggunaan alat digital dalam meningkatkan aktivitas sosial telah memberi para ilmuwan tempat data baru yang mendokumentasikan berbagai aspek kehidupan sosial (Lazer et al, 2009). Penulis bermaksud memanfaatkan data media sosial untuk mengamati dan mengukur kepribadian dari aktivitas pengguna. Media sosial yang digunakan dalam penelitian ini adalah twitter. Twitter merupakan salah satu media sosial yang digunakan di Indonesia. Pengguna mengunggah informasi yang utamanya berupa data teks pada twitter. Twitter bersifat *open data* yang memudahkan peneliti mendapatkan data tentang pengguna dengan gratis. Dibandingkan dengan situs lainnya, twitter lebih berfokus pada kecepatan dan ketangkasannya dalam menyebarkan informasi, sehingga lebih mudah bagi pengguna untuk mengunggah hal-hal yang berhubungan dengan kejadian yang terjadi secara langsung (*happening right now*) (Qui, 2012).

Terdapat banyak teori pengukuran kepribadian yang tersedia, namun penelitian ini menggunakan tes kepribadian berdasarkan teori *Big Five Personality Traits*. Teori Big Five Personality traits membagi kepribadian menjadi lima dimensi. Adapun kelima dimensi tersebut disebut dengan singkatan OCEAN, yakni *Openness to Experience*, *Conscientiousness*, *Extraversion*, *Agreeableness* dan *Neuroticism*. Teori ini menggunakan pendekatan *lexical hypothesis* sebagai pondasi utamanya, sehingga akan sesuai untuk pengukuran yang menggunakan data tekstual. *Lexical Hypothesis* secara umum dapat dijelaskan sebagai proses memahami makna terhadap data tekstual (Raad et al, 2015).

Berberapa penelitian tentang pengukuran kepribadian dengan media sosial sudah pernah dilakukan. Salah satunya adalah penelitian oleh Sarwani dan Mahmudy (2015) yang menggunakan metode klasifikasi teks dengan algoritma naïve bayes. Algoritma naïve bayes adalah algoritma yang digunakan dalam metode klasifikasi *machine learning*. Metode ini memberikan data latihan agar komputer dapat membangun model klasifikasi dari pola aturan pengklasifikasian yang ditemukan dalam data latihan.

Metode *machine learning* sangat diunggulkan penggunaannya dalam menemukan aturan pengklasifikasian pada data bervolume besar. Namun menurut Liu (2016) metode *machine learning* ini cukup sulit diterapkan terhadap pengelolaan data tekstual, karena data tekstual yang dikumpulkan seringkali hanya berupa data *noise* dan atau susunannya tidak lengkap. Hal ini akan menyebabkan komputer kesulitan untuk menemukan aturan pengklasifikasian yang diharapkan.

Penelitian lain oleh Qui (2012) menggunakan data twitter untuk pengukuran kepribadian dengan menggunakan aplikasi berbayar, yakni *Linguistic Inquiry and Word Count (LIWC)*. LIWC adalah standart analisis tekomputerisasi untuk penggunaan data linguistik/ teks. LIWC mempelajari bagaimana kata-kata yang digunakan dalam bahasa keseharian dapat

mengungkapkan pikiran, perasaan, kepribadian, motivasi dan memberikan wawasan sosial dan psikologis. Namun sejauh ini LIWC hanya mempelajari penggunaan kata dalam bahasa Inggris. Penulis belum menemukan LIWC yang menggunakan kata bahasa Indonesia.

Dalam proses menganalisis kepribadian seseorang, dibutuhkan metodologi yang sesuai dan akurat. Pengukuran kepribadian memiliki ketentuan khusus yang hanya dipahami oleh psikolog. Ontologi merupakan representasi formal mengenai spesifikasi eksplisit dari suatu kumpulan konsep. Ontologi dapat mengelompokkan kosakata dan hirarki yang memodelkan suatu domain dengan definisi objek, konsep, properti dan relasi (Noy dan McGuinness, 2000). Sesuai dengan pengertiannya, ontologi dapat digunakan untuk mengelompokkan kosakata dan hirarki yang memodelkan domain untuk menjelaskan konsep kepribadian.

Pendekatan ontologi dinilai tepat untuk pengukuran kepribadian karena pembangunan model klasifikasi ontologi berangkat dari domain pengetahuan khusus. Ontologi dapat membawa banyak keuntungan untuk pengelolaan data tekstual, yakni aturan klasifikasi dibuat jelas berdasarkan dasar domain pengetahuannya, mudah dimodifikasi jika ditemukan adanya kekurangan, dapat diintegrasikan dengan model *machine learning* dan mudah untuk dipahami (Liu, 2016). Dengan ontologi penulis akan lebih mudah membangun aturan pengklasifikasian yang sesuai dengan pengetahuan psikologi. Pengklasifikasian data tekstual yang rumit untuk metode *machine learning* dapat diatasi dengan model ontologi.

Tujuan dari penelitian ini adalah untuk membangun model ontologi sesuai teori kepribadian *Big Five Personality Traits*. Model ontologi ini menggunakan metode klasifikasi teks Bahasa Indonesia untuk mengukur kepribadian pengguna media sosial twitter. Hasil pengukuran kepribadian dapat digunakan untuk membantu keputusan sumber daya yang berkaitan dengan performansi kerja.

1.3.Perumusan Masalah

Meskipun terdapat banyak metode dan keilmuan mengenai psikologi kepribadian, namun saat ini pengukuran kepribadian dirasa cukup sulit untuk memberikan hasil yang aktual. Berdasarkan artikel yang diperoleh dari *thehrdepartment.ie* (2015), bahwa terdapat kesulitan dalam menentukan kepribadian, salah satu sebabnya adalah adanya kecenderungan untuk menjawab sesuai dengan apa yang diinginkan oleh pewawancara ketika sedang dilakukan test wawancara, serta menjawab dengan apa yang biasa dianggap baik oleh lingkungan sosial (normatif). Sehingga tes tersebut akan memberikan hasil pengukuran kepribadian yang tidak sesuai dengan kepribadian yang sebenarnya.

Adanya pemanfaatan data sosial media sebagai sumber data yang baru serta adanya permasalahan tersebut, memicu penelitian ini untuk membuat pendekatan baru dalam mengukur kepribadian seseorang. Penggunaan data media sosial yang merupakan jejak aktivitas pengguna dapat dimanfaatkan untuk membangun model klasifikasi teks untuk pengukuran kepribadian. Penulis berharap model klasifikasi tersebut dapat digunakan untuk memberikan alternatif baru dalam melakukan pengukuran kepribadian untuk kepentingan manajemen sumber daya.

1.4.Pertanyaan Penelitian

Berdasarkan perumusan masalah yang telah disebutkan sebelumnya, adapun pertanyaan penelitian dapat dijelaskan sebagai berikut:

1. Bagaimana membuat model ontologi untuk mengukur kepribadian?
2. Bagaimana membuat model ontologi berdasarkan kata-kata yang ditulis pada akun sosial media dalam bahasa Indonesia?

1.5.Tujuan Penelitian

Berdasarkan uraian pertanyaan penelitian, adapun tujuan penelitian ini adalah sebagai berikut:

1. Membangun model pendekatan klasifikasi teks berbasis ontologi untuk mengukur kepribadian seseorang berdasarkan tipe kepribadian *Big Five Personality Traits*.
2. Membangun model pendekatan klasifikasi teks dari aktivitas media sosial dalam Bahasa Indonesia dengan menggunakan metode *Ontology based Text Classification*.

1.6. Manfaat Penelitian

Manfaat penelitian ini dapat dijelaskan kedalam dua aspek, yakni aspek teori dan praktis. Berikut penjelasan manfaat penelitian dari kedua aspek:

1. Aspek Teori

Penelitian ini diharapkan mampu membuka wawasan tentang pengaplikasian metode *Text Classification* di bidang manajemen. Peneliti berharap dari penelitian ini dapat menjadi rujukan untuk penelitian selanjutnya guna memperluas pengetahuan para akademisi berkenaan dengan metode *Text Classification* menggunakan model Ontologi.

2. Aspek Praktis

Penelitian ini diharapkan dapat membantu mempermudah pengukuran kepribadian oleh manajer sumber daya manusia pada rangkaian seleksi pegawai baru. Pengukuran kepribadian dimaksudkan untuk pengambilan keputusan sumber daya yang berhubungan dengan *Job Performance*. Selain itu, penelitian ini dapat dijadikan alternatif dalam memanfaatkan data media sosial untuk menentukan keputusan manajemen.

1.7. Ruang Lingkup Penelitian

1. Lokasi dan Obyek

Penelitian ini berfokus pada *tweets* yang diunggah oleh pengguna twitter untuk pengukuran kepribadian penggunanya. Penelitian ini mengumpulkan data *timeline user* twitter di Indonesia, sehingga klasifikasi teks menggunakan Bahasa Indonesia dalam kontennya.

2. Waktu penelitian

Waktu pengambilan data dilakukan pada 25 Agustus 2017 pukul 14.00 WIB. *Tweets* yang kami dapat merupakan jejak penulisan *tweets* yang sudah diunggah pada waktu lampau.

1.7.Sistematika Penulisan Tugas Akhir

Pembahasan dari penelitian ini akan dibagi dalam lima bab yang terdiri dari beberapa sub-bab. Berikut sistematika penulisan penelitian secara garis besar:

BAB I PENDAHULUAN

Dalam BAB ini, penulis menjelaskan mengenai gambaran umum objek penelitian, latar belakang penelitian, perumusan masalah, tujuan penelitian, manfaat penelitian serta sistematika penulisan.

BAB II LANDASAN TEORI

BAB ini memuat mengenai tinjauan pustaka dan penelitian terdahulu yang berkaitan dengan penelitian.

BAB III METODE PENELITIAN

Bab ini berisikan tentang metode yang digunakan oleh peneliti, yakni berisikan tentang jenis penelitian, sumber penelitian, metode *data mining* serta proses mengolah data pada penelitian.

BAB IV HASIL PENELITIAN DAN PEMBAHASAN

Pada BAB ini penulis memaparkan hasil dan pembahasan dari penelitian yang sudah dilakukan.

BAB V KESIMPULAN DAN SARAN

Penulis menjelaskan kesimpulan dan saran yang diperoleh dari penelitian. Kesimpulan dan Saran nantinya dapat dimanfaatkan oleh objek yang bersangkutan maupun pihak-pihak yang terkait dalam penelitian.