

1. Pendahuluan

1.1 Latar Belakang

Media sosial merupakan suatu layanan *online* yang berdiri di atas suatu *platform* dan memungkinkan penggunanya untuk saling terhubung, berbagi, dan bertukar informasi. Media sosial telah menjadi bagian dari aktifitas pada beberapa kalangan, terutama kalangan remaja. Hal ini ditunjang oleh kemajuan perangkat lunak *gadget*, teknologi informasi, dan kecepatan jaringan internet yang makin pesat [1]. Terhitung pada Maret 2015, 29% dari total populasi Indonesia atau sebesar 74 juta penduduk Indonesia aktif sebagai pengguna media sosial [2]. Hingga kini, berbagai macam media sosial berkembang dengan karakteristik dan fitur yang berbeda-beda.

Instagram hadir menambah daftar panjang media sosial yang sudah ada. Instagram merupakan media sosial berbagi foto dan video [3]. Karena keunikan dari fitur yang dimiliki Instagram, saat ini Instagram menjadi media sosial yang banyak diminati oleh masyarakat. Di tahun 2013, Instagram berhasil mencapai angka 150 juta pengguna aktif setiap bulannya [3]. Di Indonesia sendiri, tahun 2015 Instagram menempati peringkat ketujuh dengan total pengguna aktif sebesar 7% dari total penduduk Indonesia [2]. Fakta tersebut menunjukkan tingginya popularitas pengguna Instagram. Hal ini mengundang pihak-pihak yang ingin mengambil keuntungan dengan cara menyebar *spam*. *Spam* secara umum diartikan sebagai pesan atau postingan yang dikirim ke sejumlah pengguna dimana pesan tersebut sebenarnya bukan permintaan dari pengguna, atau memiliki konten yang tidak sesuai [4]. Adapun *spam* pada Instagram khususnya pada komentar adalah semua komentar yang mengandung iklan/penawaran produk atau jasa. Dengan banyaknya responden yang terlibat maka akan semakin meningkatkan peluang keuntungan yang bisa didapat.

Mengatasi permasalahan ini, Instagram akhirnya merilis fitur untuk menyaring komentar-komentar yang tidak diinginkan pengguna. Fitur ini bekerja secara manual, dimana pengguna memasukkan kata kunci yang ingin diblokir pada sebuah kolom. Namun fitur ini masih memiliki kekurangan, yaitu untuk setiap kata kunci yang diinputkan memiliki kemungkinan muncul pada komentar yang sebenarnya tidak mengganggu.

Adanya masalah di atas menjadi dasar penelitian ini. Pada penelitian ini dirancang fitur pendeteksi *spam* yang dapat membedakan konteks komentar yang mengganggu dengan konteks komentar yang tidak mengganggu dengan memanfaatkan ontologi. Ontologi menyediakan skema yang tepat untuk sumber konseptual yang luas [5]. Sumber konseptual yang dimaksud adalah komentar pada Instagram yang saat ini memiliki *range* yang luas dan beranekaragam. Melalui skema tersebut, ontologi akan menjadi kamus yang berisikan istilah dari definisi *spam* yang telah ditentukan serta menyediakan spesifikasi terhadap istilah-istilah yang dimodelkan.

1.2 Perumusan Masalah

Berikut ini merupakan masalah yang diangkat pada penelitian ini.

1. Bagaimana mengidentifikasi *spam* komentar pada Instagram menggunakan ontologi?
2. Bagaimana perfomansi dari hasil identifikasi *spam* menggunakan ontologi?

Adapun batasan masalah dari penelitian ini dijabarkan sebagai berikut.

1. *Spam* berfokus pada bagian komentar.
2. Yang dimaksud dengan *spam* komentar adalah semua komentar yang merupakan iklan/penawaran barang atau jasa
3. Komentar tidak terkait *caption* pada gambar.

1.3 Tujuan

1. Untuk mengimplementasikan ontologi dalam mengidentifikasi *spam* komentar pada Instagram
2. Untuk mengukur performansi dari hasil identifikasi *spam* pada Instagram menggunakan ontologi

1.4 Metodologi Penyelesaian Masalah

1. Identifikasi Permasalahan

Merupakan tahap identifikasi masalah yang menjadi landasan dilaksanakannya penelitian. Pada tahap ini dilakukan pengumpulan informasi dan melakukan studi literatur terkait perancangan ontologi yang didapat dari jurnal, *paper*, dan penelitian yang sudah pernah dilakukan.

2. Identifikasi Kebutuhan

Merupakan tahap dilakukannya identifikasi kebutuhan untuk penelitian dan kebutuhan pada sistem. Kebutuhan penelitian dipenuhi dengan cara menyebar kuesioner untuk mendukung definisi *spam* pada Instagram. Teknik yang digunakan adalah teknik sampel, yaitu penelitian yang hanya melibatkan 30 orang dari total populasi. Sedangkan kebutuhan sistem terdiri dari fungsionalitas yang diperlukan serta spesifikasi perangkat keras dan lunak yang mendukung penelitian.

3. Perancangan Sistem

Tahap untuk melakukan perancangan sistem ontologi berdasarkan pada hasil pengumpulan data dan kebutuhan sistem yang telah ditentukan.

4. Implementasi Sistem

Tahap penerapan metode yang telah ditentukan untuk mengidentifikasi *spam* komentar, yaitu ontologi. Bahasa yang telah direkomendasikan yaitu OWL (*Ontology Web Language*). Pada tahap ini juga dilakukan pembangunan sistem *spam*

detector untuk melakukan klasifikasi berdasarkan pada dataset yang telah melalui proses *preprocessing*. Untuk lebih jelasnya implementasi sistem akan dijabarkan pada bab 3.

5. Pengujian

Merupakan tahap menguji sistem yang telah dibangun. Pengujian bertujuan untuk mengetahui sistem dapat memberikan output yang sesuai dengan kebutuhan penelitian serta dapat menyelesaikan permasalahan yang melatarbelakangi penelitian ini.

6. Analisis Hasil Pengujian

Tahap menganalisis hasil pengujian sistem dilihat dari performansi yang dihasilkan oleh sistem. Performansi sistem ditunjukkan dengan nilai *precision*, *recall*, dan *f1-score*.

1.5 Sistematika Penulisan

Tugas akhir dengan judul “**Identifikasi Spam Komentar pada Instagram Berbasis Ontologi**” ini disusun dengan sistematika penulisan sebagai berikut :

1. BAB 1 – Pendahuluan

Bab 1 menjelaskan latar belakang dilakukannya penelitian, perumusan masalah serta batasan masalahnya, tujuan penelitian, metodologi penyelesaian masalah serta sistematika penulisan dari tugas akhir ini.

2. BAB 2 – Kajian Pustaka

Bab 2 menjelaskan teori-teori yang mendasari dan mendukung pengerjaan tugas akhir.

3. BAB 3 – Perancangan Sistem

Bab 3 menjabarkan proses perancangan dari sistem yang akan dibangun, dimulai dari identifikasi masalah, identifikasi kebutuhan, proses pengumpulan dataset, tahap-tahap membangun ontologi serta tahap-tahap melakukan klasifikasi pada sistem.

4. BAB 4 – Pengujian dan Analisis

Bab 4 menjelaskan dataset, ontologi, dan kamus yang digunakan, tujuan dilaksanakan pengujian, skenario yang digunakan selama pengujian serta analisis dari hasil pengujian.

5. BAB 5 – Kesimpulan dan Saran

Bab 5 berisi kesimpulan hasil dari penelitian yang telah dilakukan serta memuat saran-saran yang dapat membangun penelitian selanjutnya.