

# 1. PENDAHULUAN

## 1.1. Latar Belakang

YouTube merupakan situs web sharing video nomor satu di dunia dengan total lalu lintas pengunjung daring sebesar 1,7 milyar setiap bulannya (25 Juni 2019). Belum lama ini YouTube Indonesia sedang ramai oleh video-video pelantikan calon-calon menteri Kabinet Indonesia Maju Jokowi-Ma'ruf. Situs berita daring berlomba-lomba untuk mengunggah momen pelantikan tersebut. Pro-kontra pun bermunculan di kolom komentar. Dari sekian banyak video pelantikan di YouTube, situs berita daring Kompas, CNN Indonesia, dan SCTV, dengan nama kanal masing-masing KOMPASTV, CNN Indonesia, dan Surya Citra Televisi (SCTV), mendapatkan jumlah view paling banyak dibandingkan dengan yang lainnya. Begitupun juga dengan komentar dari video yang diunggah mencapai jumlah masing-masing 1,3 juta, 693 ribu, dan 1,2 juta. Maka dari itu akan sangat sulit untuk memahami apakah warganet pengguna situs YouTube cenderung pro, kontra, atau netral dengan hasil calon menteri yang diusung oleh pemerintahan Jokowi-Ma'ruf. Dibutuhkan teknik untuk menganalisis komentar yang begitu banyak supaya komentar-komentar tersebut dapat berguna dan dapat ditarik suatu kesimpulan yang bermakna.

Analisis sentimen adalah teknik untuk menganalisis opini, sentimen, penilaian dan emosi terhadap suatu entitas seperti produk, jasa, kejadian atau atribut lainnya [1]. Pemikiran dasar dari teknik analisis sentimen adalah untuk mengelompokkan teks, kalimat, atau dokumen kemudian menentukan teks, kalimat, atau dokumen tersebut termasuk ke dalam sentimen atau opini yang positif, negatif, atau netral [2]. Metode klasifikasi yang digunakan bisa menggunakan 2 cara yaitu lexicon based dan machine learning based. Metode lexicon based menggunakan kamus, kosa kata, atau korpus yang sudah ada (yang telah dibuat oleh para ahli linguistik) dan mencocokkannya dengan data yang ingin diklasifikasikan. Sedangkan metode machine learning menggunakan algoritma machine learning untuk mengklasifikasikan data yang ingin diklasifikasikan [3]. Terdapat korpus leksikal yang cukup terkenal yaitu SentiWordNet. SentiWordNet merupakan hasil anotasi otomatis dari semua synset WordNet [4]. SentiWordNet sampai saat ini belum mendukung bahasa Indonesia, maka dari itu penulis juga menggunakan sentiment lexicon Indonesia [5] sebagai kamus leksikon pada tahap ekstraksi fitur dan akan

membandingkannya dengan lexical resource SentiWordNet.

Pada penelitian tugas akhir ini akan dilakukan analisis sentimen pada komentar di situs web YouTube dengan domain pelantikan calon menteri Kabinet Indonesia Maju Jokowi-Ma'ruf. Terdapat 3 kanal besar yang akan dianalisis komentar videonya yaitu KOMPASTV dengan judul video "Resmi! Daftar Nama Menteri Kabinet Indonesia Maju Jokowi-Ma'ruf", CNN Indonesia dengan judul video "Sah! Presiden Jokowi Umumkan Menteri Baru Kabinet Indonesia Maju", dan SCTV dengan judul video "Ini Nama-nama Menteri Baru Indonesia Maju Periode 2019-2024 – Pengumuman Kabinet Jokowi-Ma'ruf". Untuk metode pengklasifikasian menggunakan lexicon based dengan membandingkan 2 tahapan ekstraksi fitur yang berbeda yaitu menggunakan korpus leksikal SentiWordnet dan sentiment lexicon Indonesia.

## **1.2. Perumusan Masalah**

Berdasarkan latar belakang diatas, maka rumusan masalah yang akan dibahas pada penelitian ini adalah.

1. Bagaimana hasil sentimen dari komentar warganet situs YouTube terhadap terpilihnya menteri kabinet Indonesia Maju?
2. Bagaimana hasil perbandingan performansi sistem dari percobaan ekstraksi fitur menggunakan SentiWordNet dan sentiment lexicon Indonesia?

## **1.3. Tujuan**

Tujuan pembuatan Tugas Akhir ini adalah.

1. Mengetahui hasil sentimen dari komentar warganet situs YouTube terhadap terpilihnya menteri kabinet Indonesia Maju
2. Mengetahui hasil perbandingan performansi sistem dari percobaan ekstraksi fitur menggunakan SentiWordNet dan sentiment lexicon Indonesia

## 1.4. Batasan Masalah

Batasan pembuatan Tugas Akhir ini adalah.

1. Data komentar yang diambil menggunakan bahasa Indonesia.
2. Data komentar hanya diambil dari 3 kanal berita terbesar di Indonesia yaitu CNN, Kompas TV, dan SCTV.
3. Untuk mengetahui performansi sistem yang akan dibangun, penulis membuat data komentar yang telah dilabeli manual.
4. Berfokus pada hasil klasifikasi sentimen dan perbandingan performansi sistem antara SentiWordNet dan sentiment lexicon Indonesia.

## 1.5. Rencana Kegiatan

Untuk menyelesaikan permasalahan yang diangkat, perlu dilakukan tahap-tahap penyelesaian masalah, berikut adalah tahapan yang digunakan:

### 1. Studi Literatur

Pada tahap studi literatur dilakukan untuk pengumpulan paper, jurnal, artikel dan buku sebagai sumber referensi untuk menunjang penelitian mengenai analisis sentiment.

### 2. Analisis

Hal-hal yang dilakukan dalam tahap ini adalah menganalisis tahap demi tahap dari sentimen analisis dan cara kerja dari ekstraksi fitur menggunakan SentiWordNet dan sentiment lexicon Indonesia.

### 3. Pengumpulan Data

Pada tahap pengumpulan data dilakukan pengumpulan data komentar YouTube menggunakan *YouTube Comment Crawler*.

#### 4. Implementasi Sistem

Pada tahap ini akan dilakukan proses implementasi pengkodean program dalam aplikasi komputer menggunakan bahasa pemrograman python.

#### 5. Pengujian Sistem dan Analisis Hasil

Pada tahap ini dilakukan proses pengujian dan percobaan terhadap sistem sesuai dengan spesifikasi yang ditentukan sebelumnya serta memastikan program yang dibuat dapat berjalan seperti yang diharapkan.

#### 6. Penyusunan Laporan

Pada tahap ini dilakukan penulisan dokumentasi hasil analisis dan implementasi. Analisis Kebutuhan dan Perancangan Sistem

### 1.6. Jadwal Kegiatan

Tabel 1.1 Jadwal Kegiatan Semester Ganjil Tahun Ajaran 2019/2020

Kegiatan	Bulan			
	September	Oktober	november	Desember
Studi Literatur				
Analisis				
Pengumpulan Data				

Tabel 1.2 Jadwal Kegiatan Semester Genap Tahun Ajaran 2019/2020

Kegiatan	Bulan					
	Januari	Februari	Maret	April	Mei	Juni
Studi Literatur						
Analisis						
Pengumpulan Data						
Implementasi Sistem						

Pengujian Sistem dan Analisis Hasil						
Penyusunan Laporan						

## 1.7. Sistematika Penulisan Tugas Akhir

Sistematika penulisan tugas akhir ini dibagi menjadi lima bab yaitu :

### **BAB I PENDAHULUAN**

Pada bab I akan membahas mengenai latar belakang, perumusan masalah, tujuan, Batasan masalah, metodologi penelitian dan sistematika penulisan.

### **BAB II TINJAUAN PUSTAKA DAN LINGKUP PENELITIAN**

Dalam bab II berisi mengenai pemaparan landasan teori dan konsep sebagai dasar untuk menganalisa permasalahan yang berkaitan dengan penelitian.

### **BAB III METODE PENELITIAN**

Pada bab III ini berisi tentang perancangan dari sistem dan penguraian jenis penelitian

### **BAB IV ANALISIS DAN PEMBAHASAN**

Dalam bab IV menjelaskan dan menguraikan secara sistematis sesuai dengan perumusan masalah dan tujuan masalah penelitian ini. Menjelaskan mengenai hasil pengujian sistem dan analisis terhadap hasil pengujian dan perbandingan, meliputi metode, tingkat akurasi sistem, persisi, kualitas dan kesalahan.

### **BAB V KESIMPULAN DAN SARAN**

Pada bab V berisi kesimpulan yang diperoleh dari hasil penelitian dan analisis pada bab sebelumnya. Pada bab ini dilengkapi juga saran-saran yang dapat dilakukan untuk mengembangkan penelitian ini lebih lanjut.