

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Saat ini internet telah menyentuh berbagai aspek, peran internet dan data mining sangat besar dalam proses pencarian dan pengolahan informasi [1]. Salah satu aspek yang tersentuh oleh internet adalah sosial masyarakat di situs sosial media online. Namun tidak dapat dimungkiri dari sekian banyaknya opini tidak selalu sama yaitu netral (fakta, ucapan selamat, dll), pertanyaan, sentiment positif dan sentiment negatif [2]. Dalam kasus yang erat kaitannya Penerapan Pembatasan Sosial Berskala Besar (PSBB) diyakini merupakan cara paling ampuh untuk menekankan peserta didik tentang penularan pandemi Covid-19[3]. Dalam pendidikan, kebijakan ini melarang proses belajar mengajar secara tatap muka dan berubah menjadi daring (dalam jaringan).

Hal ini banyak menimbulkan masalah-masalah seperti ketidaksiapan teknologi, psikologis peserta didik, dan banyak peserta didik merasa kurang nyaman sehingga menuai berbagai komentar melalui media sosial. Sosial media twitter digunakan sebagai salah satu media yang dapat menunjukkan ekspresi terhadap respon peserta didik dalam kegiatan pendidikan yang berlangsung selama pandemi Covid-19 di Indonesia[4].

Untuk mengetahui opini peserta didik lewat tweetnya mengenai senang atau tidaknya terhadap pembelajaran daring ini maka diperlukan suatu sistem analisis sentiment, dan juga untuk mengklasifikasi dokumen yang mana merupakan pemberian kategori, kepada dokumen yang belum memiliki kategori dan memiliki ciri unik yang sama [30]. Analisis sentimen disebut tambang pendapat, melibatkan dalam membangun sistem untuk mengumpulkan dan meneliti pendapat[5]. Tetapi dengan menggunakan analisis sentiment saja memiliki kekurangan diantaranya hanya berada pada level dokumen atau kalimat saja. Performa yang dihasilkan jauh berbeda ketika menggunakan analisis sentimen pada level aspek. Dengan basis aspek, penelitian ini dapat menentukan apakah suatu sentimen bernilai positif atau negatif kesesuaian terhadap aspek yang berelasi pada setiap kalimat. Sehingga nantinya dengan sistem ini, dapat menjawab kebijakan belajar daring apakah sudah sesuai penerapannya, atau masih ada kekurangan yang harus diperbaiki berdasarkan opini masyarakat dan aspeknya dapat membantu sektor mana saja yang harus dibenahi.

Dari beberapa penelitian terdahulu diantaranya yang pernah dilakukan oleh Salsabila Rahma Yustihan dkk, tentang Analisis Sentimen Basis Aspek Untuk Ulasan Rumah Makan menggunakan SVM menggunakan komposisi sentimen Positif sebanyak 285 data latih dan

data uji sebanyak 103 dan Negatif data latih ada 75 dan data uji sebanyak 28 data. Dengan akurasi 86%, presisi 86% dan F1-Score 92% [6]. Di penelitian serupa dari Wirdhayanti Paulina dkk pada Analisis Sentimen Berbasis Aspek Ulasan Pelanggan Terhadap Kertanegara Premium Guest House menggunakan SVM, menghasilkan akurasi untuk sentimen aspek kamar 79%, harga 70%, dan layanan di 83% yang masing-masing sama di sentimen positif dan negatif [7]. Sedangkan di penelitian lain sebagai acuan penelitian ini dalam sentimen berbasis aspek menggunakan SVM dilakukan oleh Sabrah Ailiyya [8]. Mereka melakukan analisis sentimen berbasis aspek pada ulasan aplikasi Tokopedia dengan menggunakan *Support Vector Machine* (SVM), Hasil klasifikasi sentimen dan aspek berturut-turut menunjukkan akurasi sebesar 69,6% dan 74,2%.

Beberapa penelitian telah menggunakan SVM pada berbagai penerapan, diantaranya adalah pada pengenalan citra, analisis medik, ataupun untuk melakukan prediksi secara spesifik. Pada beberapa penelitian dijelaskan bahwa SVM adalah metode yang efisien [9]. Namun dari beberapa penelitian tersebut, belum menggunakan implementasi SVM dengan ekstraksi fitur menggunakan teknik TF-IDF dalam kaitan aspek dan sentimen. Berkaitan dengan penelitian yang telah ada dan masalah yang ditemukan maka penulis tertarik untuk melakukan penelitian analisis sentimen berbasis aspek pada opini peserta didik pengguna Twitter dengan menerapkan metode klasifikasi *Support Vector Machine* (SVM) dan *Term frequenc-Inverse document frequency* (TF-IDF) sebagai Term Weighting. Maka dari itu judul penelitian ini adalah “Analisis Sentimen Berbasis Aspek Terhadap Pembelajaran Daring Di Masa Pandemi Covid-19 Menggunakan *Support Vector Machine* (SVM)”.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka diperoleh sebuah perumusan masalah yang ada yaitu bagaimana pengaruh dan hasil analisis sentimen berbasis aspek menggunakan *Support Vector Machine* (SVM) serta TF-IDF sebagai term-weighting dalam kasus Pembelajaran Daring di Masa Pandemi Covid-19.

1.3 Maksud dan Tujuan

Berdasarkan permasalahan yang ada, maksud dari penelitian ini adalah untuk menerapkan metode *Support Vector Machine* (SVM) beserta TF-IDF sebagai term-weighting dalam klasifikasi analisis sentimen berbasis aspek pengguna *Twitter* terhadap Pembelajaran Daring. Adapun tujuan dari penelitian ini, yaitu untuk mengetahui hasil dari penerapan metode *Support Vector Machine* (SVM) jika menggunakan TF-IDF untuk term-weight dalam suatu sistem analisis sentimen berbasis aspek.

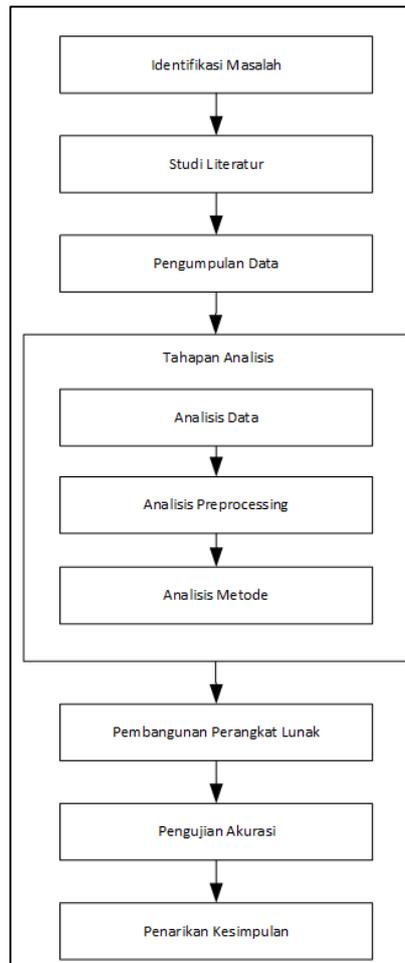
1.4 Batasan Masalah

Agar penelitian yang dilakukan semakin sesuai dengan tujuan yang diinginkan, maka batasan masalah dalam penelitian ini yaitu sebagai berikut :

1. Sumber dataset berupa opini yang berkaitan dengan Pembelajaran Daring dari mesin pencarian di *Twitter* dan data tweet berupa bahasa Indonesia.
2. Data tweet diambil dari Twitter dengan menggunakan kata kunci “Kuliah Online”.
3. Sistem yang akan dibangun akan melakukan klasifikasi sentimen pada setiap aspek. Beberapa aspek yang akan diklasifikasikan yaitu Ekonomi, Teknologi, Pendidikan, Kesehatan dengan menggunakan SVM multiclass menggunakan yang paling umum yaitu One-to-Rest.
4. Klasifikasi sentimen yang akan dibuat terdiri dari kelas sentimen positif dan negatif saja untuk tahapan opini/tweet.
5. Format *file* yang akan digunakan pada data latih dan uji adalah *Comma Separative Value* (.CSV).
6. Data diambil dengan cara *Crawling* menggunakan API (*Application*

1.5 Metodologi Penelitian

Dalam penelitian ini menggunakan pendekatan kuantitatif, salah satu karakteristik pendekatan kuantitatif adalah dapat melakukan generalisasi dengan melihat fenomena atau kejadian-kejadian sosial yang terjadi. Pendekatan kuantitatif berasal dari ilmu pengetahuan dan dijabarkan dalam bentuk *experiment reality* ataupun *aggreement reality*, dimana penelitian ini akan fokus pada eksperimen *reality*[10]. Dalam menyelesaikan penelitian ini, tentunya perlu basis yang benar agar mencapai tujuan penelitian yang sudah dipaparkan sebelumnya, dimana keterkaitannya adalah terlibatnya sumber ilmu pengetahuan, data yang representatif (tentu data yang bersifat kuantitatif), dan ada beberapa proses analisis data serta penerapan metode pengujian. Ini didukung oleh Sugiyono[11], metode kuantitatif digunakan untuk meneliti populasi atau sample tertentu, teknik pengambilan sample pada umumnya dilakukan secara random. Dimana ini sangat representatif untuk topik serta penggunaan data pada analisis sentiment. Berikut tahapan Alur penelitian yang dilakukan seperti pada Gambar 1.1.



Gambar 1.1 Alur Penelitian

Berikut penjelasan mengenai setiap langkah pada metode penelitian yang telah dipaparkan sesuai gambar 1.1 di atas yaitu:

1. Identifikasi Masalah

Dalam setiap penelitian yang dilakukan mesti didasari oleh suatu problem ataupun masalah yang signifikan, sehingga dengan menyelesaikan masalah yang diperoleh dari sebuah penelitian dapat mencerminkan keberhasilan dari penelitian tersebut. Tahapan awal dari sebuah metode penelitian adalah melakukan identifikasi fakta dan kejadian yang terjadi saat ini dan melihat sebuah problem yang akan diangkat dan dibahas, maka setelah itu dapat dilakukan sebuah pencarian solusi yang tentunya akan melalui beberapa proses lagi hingga masalah dapat terjawab secara proporsional dan jelas. Dalam kasus ini, problem yang harus diidentifikasi adalah tentang masalah analisis sentimen serta kaitannya dengan level aspek dan penerapan dari metode *Support Vector Machine (SVM)*.

2. Studi Literatur

Setelah melakukan identifikasi masalah, langkah berikutnya adalah melakukan beberapa proses dimana disebutkan sebagai proses awal dari penyelesaian masalah yang telah ditemukan, yaitu proses pengumpulan data yang berkaitan dengan masalah yang telah ditemukan tadi, baik itu berupa studi atau ilmu-ilmu terkait dan representasi data (kuantitatif) yang berkaitan dengan penelitian yang akan dijalankan, sehingga nanti pada proses selanjutnya data ini akan digunakan dan akan berubah menjadi suatu informasi yang dapat menyelesaikan masalah yang ditemukan sebelumnya. Langkah berikutnya yang dilakukan salah satunya adalah studi literatur dimana pengumpulan data yang dilakukan dengan cara mengumpulkan literatur, jurnal, paper, dan buku ataupun informasi terkait yang berkaitan dengan analisis sentimen maupun analisis sentimen berbasis aspek serta komponen pendukung dari penelitian ini.

3. Pengumpulan Dataset

Setelah literatur dan pendukung lainnya sudah dirasa cukup memadai, maka langkah berikutnya adalah menyiapkan representasi data yang akan digunakan dalam penelitian, untuk itu dalam penelitian yang kaitanya dengan analisis sentiment, biasanya data disebut dengan “dataset”. Maka dataset ini harus dikumpulkan secara kolektif dari sumber yang telah ditentukan dan sesuai dengan data yang dibutuhkan untuk melakukan analisis sentiment. Untuk itu data yang digunakan pada penelitian ini berupa opini terhadap sistem pembelajaran daring yang diambil dari situs sosial media *Twitter* dengan cara crawling dengan menggunakan API (*Application Programming Interface*).

4. Tahapan Analisis

Langkah berikutnya setelah kita mengumpulkan data sesuai pada tahapan metode penelitian pada gambar 1.1 yaitu adalah tahapan Analisis sedangkan secara rinci tahapan analisis terdiri dari beberapa langkah yaitu:

- a) Analisis Data, pada tahap ini yaitu menganalisis beberapa kelompok atau kalimat pada data yang ada, data yang dianalisis adalah kelompok atau kalimat berupa Tweet pada Twitter mengenai Sistem Pembelajaran Daring di Masa Pandemi Covid-19.
- b) Analisis Preprocessing, pada tahap ini kalimat yang terkumpul akan diproses sehingga menjadi data yang terstruktur dan mudah untuk diolah. Langkah-langkah Preprocessing yang akan dilakukan terdiri dari *Case Folding*, *Remove Punctuation*, *Normalisasi Fitur*, *Tokenizing*, *Stopwords Removal*, dan

Stemming.

- c) Analisis Metode, pada tahap ini akan dilakukan analisis pada proses penggunaan metode klasifikasi atau proses pembentukan model klasifikasi. Untuk penelitian ini akan menggunakan metode *Support Vector Machine (SVM)* dengan menggunakan tambahan proses sebelumnya meliputi pembobotan kata dengan menggunakan *Term Frequenc-Inverse Document Frequency (TF-IDF)*.

5. Pembangunan Perangkat Lunak

Setelah semua proses analisis dilakukan selanjutnya membangun sebuah perangkat lunak yang tentunya berbasis tatap muka/interface. Serta tentunya proses pembangunan ini diluar dari proses dalam membangun analisis yang menggunakan Python. Untuk analisis sentiment ini akan dibangun sebuah antarmuka berbasis web sederhana, yang hanya menampilkan hasil dari analisis sentimen.

6. Pengujian Akurasi

Setelah tahapan analisis selesai dilakukan, model klasifikasi yang menggunakan SVM akan berhasil merepresentasikan data training yang digunakan dalam proses *learning*/pembelajaran maka tahapan berikutnya melakukan proses pengujian, dimana dalam pengujian ini yang akan dilakukan adalah memasukkan data uji/testing ke dalam model yang dimiliki kemudian hasil dari pengujian berupa komponen-komponen seperti akurasi, *recall* dan *f1 measure* akan dilihat seperti apa, apakah representatif atau tidak. Serta dalam melakukan proses pengujian akan menggunakan *cross-validation*, yang akan dijelaskan dalam bab berikutnya secara terperinci.

7. Penarikan Kesimpulan

Langkah terakhir dalam metode penelitian ini adalah penarikan kesimpulan, dimana penarikan kesimpulan akan berdasarkan dari hasil yang diperoleh dari beberapa proses terutama pada hasil analisis dan pengujian serta melihat korelasi dari masalah awal yang ada yang nantinya akan menjadi sebuah kesimpulan apakah masalah terjawab secara signifikan atau tidak.

1.6 Sistematika Penulisan

Sistematika penulisan yang digunakan dalam laporan hasil penelitian ini adalah cara pembagian pembahasan laporan menjadi beberapa bab dan sub-bab. Pembagian bab secara umum adalah sebagai berikut:

BAB 1 PENDAHULUAN

Bab ini akan membahas kerangka penelitian atau percobaan dalam penelitian, meliputi latar belakang permasalahan, perumusan masalah, menentukan maksud dan tujuan penelitian, batasan masalah, metode penyelesaian masalah serta sistematika penulisan dari penelitian Analisis Sentimen Berbasis Aspek Terhadap Pembelajaran Daring Di Masa Pandemi Covid-19 Menggunakan *Support Vector Machine (SVM)*

BAB 2 TINJAUAN PUSTAKA

Bab ini memuat berbagai konsep dasar dan teori-teori yang berkaitan dengan penelitian yang akan dilakukan, diantaranya mengenai konsep penelitian-penelitian terdahulu seperti analisis sentimen, analisis sentimen berbasis aspek, teknik preprocessing, ekstraksi fitur, metode SVM, pemodelan, dan sistem yang akan digunakan.

BAB 3 ANALISIS SISTEM

Bab ini berisi tentang tahapan analisis dan perancangan, Analisis yang akan dibahas berupa analisis masalah, analisis data, gambaran umum system, analisis proses, analisis pengujian, kebutuhan perangkat lunak, kebutuhan non-fungsional, kebutuhan fungsional, dan perancangan antar muka.

BAB 4 IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab ini berisi tentang implementasi dan pengujian dari perangkat lunak yang dibangun sesuai dengan analisis dan perancangan yang telah dibuat. Mulai dari implementasi hardware dan software yang digunakan, implementasi antarmuka, serta hasil pengujian. Hasil pengujian yang dilakukan yaitu mengukur tingkat akurasi.

BAB 5 KESIMPULAN DAN SARAN

Pada bab ini akan diuraikan kesimpulan dari penelitian yang telah dilakukan, serta akan membahas tentang saran yang dapat dijadikan masukan untuk penelitian selanjutnya.