

BAB 1

PENDAHULUAN

1.1. Latar Belakang Masalah

Analisis sentimen atau *opinion mining* merupakan proses memahami, mengekstrak, dan mengolah data tekstual secara otomatis untuk mendapatkan suatu informasi sentimen yang terkandung dalam suatu kalimat opini [1]. *Aspect Based Sentiment Analysis* (ABSA) adalah salah satu domain kasus *opinion mining* yang bertujuan untuk mendeteksi polaritas teks tertulis berdasarkan dengan aspek tertentu. Klasifikasi sentimen terhadap aspek dengan menentukan opini terhadap suatu aspek dari suatu entitas ke dalam kategori positif, negatif atau netral, atau menggunakan angka dalam bentuk rating terhadap aspek [2]. *Learning Vector Quantization* (LVQ) adalah salah satu metode pembelajaran pada lapisan kompetitif jaringan saraf tiruan yang akan secara otomatis belajar untuk mengklasifikasi inputan ke dalam kelas tertentu [3].

Analisis sentimen merupakan salah satu masalah penelitian yang populer baik dalam aplikasi kehidupan nyata atau dalam bidang akademik. Penelitian sebelumnya yang terkait dengan menggunakan LVQ yang telah dilakukan untuk mengklasifikasi dokumen dan menganalisa sentimen diantaranya, analisis sentimen dengan menggunakan Metode LVQ [4] yang menghasilkan akurasi 71,15%. Kemudian LVQ untuk Klasifikasi abstrak tesis [5] menghasilkan akurasi 90%. Sedangkan pada penelitian [7] di bidang klasifikasi teks dengan komparasi algoritma LVQ, RBF, KNN, Naïve Bayes, dan C4.5, algoritma LVQ menghasilkan akurasi sebesar 84%. Oleh karena itu, algoritma LVQ memiliki kemungkinan untuk memperoleh akurasi yang lebih baik apabila dilakukan pada bidang klasifikasi analisis sentimen.

Berdasarkan Penelitian tersebut, dapat disimpulkan metode LVQ dapat digunakan untuk kasus Analisis Sentimen dan dapat menghasilkan akurasi yang baik dibandingkan metode lainnya. Dalam kasus analisis sentimen menggunakan LVQ hanya ada beberapa penelitian yang telah dilakukan dan penelitian tersebut belum berdasarkan aspek. Berdasarkan hal tersebut, maka penelitian ini dilakukan untuk menerapkan LVQ pada kasus analisis sentimen berdasarkan aspek dengan dataset bahasa indonesia.

1.2. Identifikasi Masalah

Berdasarkan uraian latar belakang diatas, maka didapat permasalahan yaitu rendahnya tingkat akurasi yang diperoleh dari penelitian sebelumnya dalam kasus analisis sentimen.

1.3. Maksud dan Tujuan

Maksud dari penelitian ini adalah membangun sistem analisis sentimen berdasarkan aspek dengan metode *LVQ*. Adapun tujuan dari penelitian ini adalah mengukur akurasi metode *LVQ* untuk kasus sentimen berdasarkan aspek pada dataset bahasa Indonesia.

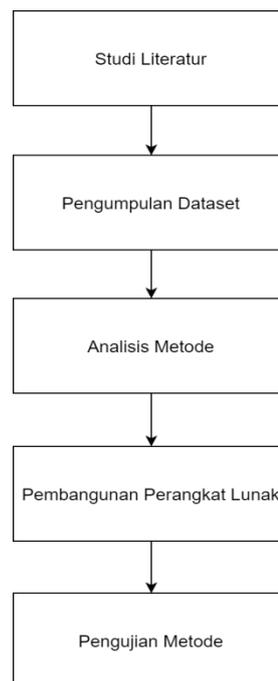
1.4. Batasan Masalah

Agar penelitian ini dapat dilakukan lebih fokus dan terarah, maka dibuat batasan masalah sebagai berikut:

1. Data Masukan
 - a. Format file dataset adalah .csv.
 - b. Data yang digunakan adalah dataset yang berasal dari penelitian sebelumnya. [8]
 - c. Teks yang diklasifikasikan dalam Bahasa Indonesia.
 - d. Aspek sudah ditentukan yang berupa aspek makanan, harga, pelayanan, dan tempat.
2. Proses
 - a. *Preprocessing* yang digunakan yaitu *casefolding*, *filtering*, *tokenization*, *word normalization* dan *stopword removal*.
 - b. Menggunakan *TF-IDF* untuk pembobotan kata.
3. Data Keluaran
 - a. Data yang dikeluarkan berupa Hasil Akurasi yang akan disimpan dalam file berformat (.txt).

1.5. Metodologi Penelitian

Metodologi penelitian yang digunakan dalam penelitian ini adalah metode kuantitatif [9]. Metode ini digunakan karena data yang digunakan dalam penelitian ini berbentuk angka dan bersifat fakta serta bisa diukur secara akurat dengan alat yang objektif. Penelitian ini dilakukan berdasarkan prosedur yang telah direncanakan sebelumnya sehingga terdiri dari tahap-tahap kegiatan. Adapun tahap penelitian yang akan dilakukan dengan alur seperti pada Gambar 1.1



Gambar 1.1 Alur Penelitian [9]

Adapun penjelasan langkah-langkah yang dilakukan dalam penelitian ini adalah sebagai berikut.

1.5.1. Studi Literatur

Studi ini dilakukan dengan cara mempelajari, meneliti dan menelaah berbagai literatur yang bersumber dari buku-buku, teks, jurnal ilmiah, situs-situs di internet, dan bacaan yang terkait dengan topik analisis sentimen berdasarkan aspek, metode preprocessing, metode *TF-IDF*, dan metode klasifikasi *LVQ*.

1.5.2. Pengumpulan Dataset

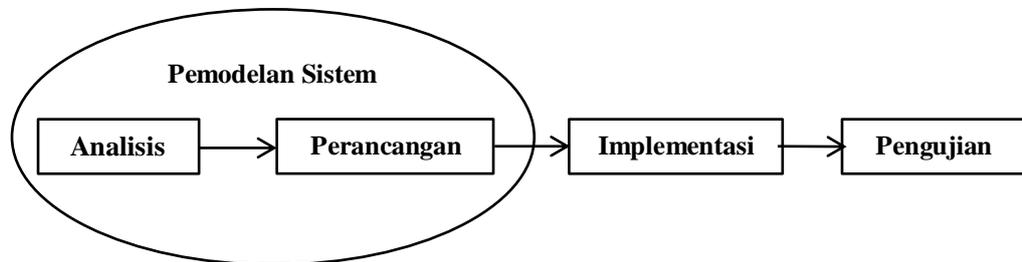
Data yang akan dikumpulkan disini terbagi menjadi dua, yaitu dataset untuk latih dan dataset untuk data uji. Data latih dan data uji yaitu berisi data Review dari penelitian sebelumnya [8].

1.5.3. Analisis Metode

Selanjutnya setelah dataset terkumpul masuk ke dalam tahap analisa, dimana tahap ini data literatur dan referensi yang sudah terkumpul dipelajari terkait metode yang akan digunakan yaitu *LVQ*, dan fitur apa saja yang dibutuhkan dan hal lainnya terkait pembangunan sistem termasuk menganalisa bagaimana sistem akan berjalan mulai dari tahap *input* kemudian proses hingga hasil berupa teks hasil pengenalan sistem.

1.5.4. Pembangunan Perangkat Lunak

Model yang akan digunakan pada penelitian ini adalah model *Sequential Linear Model* [9] karena tahapan dalam penelitian ini bersifat sekuensial mulai dari pengumpulan data, tahap *preprocessing*, klasifikasi, hingga pengujian dilakukan secara berurutan. Model pembangunan perangkat lunak *Sequential Linear Model* bisa dilihat pada Gambar 1.2.



Gambar 1.2 Model *Sequential Linear Model*

Berikut adalah tahap - tahap pada *Sequential Linear Model*:

1. Analisis, menganalisis segala kebutuhan yang diperlukan dalam pembuatan sistem, seperti kebutuhan perangkat lunak dan perangkat keras, dan lain-lain.
2. Desain sistem dan perangkat lunak, tahap ini meliputi pembuatan rancangan sistem analisis sentimen berdasarkan aspek.
3. Implementasi, tahap pembuatan aplikasi sesuai dengan desain yang ditentukan pada tahap perancangan. Tahap ini meliputi *coding* dan implementasi algoritma *LVQ*.

4. Pengujian, dilakukan pengujian terhadap sistem yang telah dibuat untuk mengetahui seberapa besar akurasi yang dihasilkan oleh metode yang diterapkan di kasus analisis sentimen berdasarkan aspek.

1.5.5. Pengujian

Pengujian sistem, perangkat lunak yang telah dibangun akan dievaluasi jika terdapat kekurangan atau kesalahan kode program. Selain itu, akan dilakukan pengujian terhadap fungsionalitas sistem dengan menggunakan *Black-Box*, serta akan dilakukan juga pengujian untuk mengukur tingkat akurasi yang dihasilkan metode *LVQ* untuk kasus analisis sentimen berdasarkan aspek.

1.6. Sistematika Penulisan

Sistematika penulisan disusun untuk memberikan gambaran secara umum mengenai bahasan penelitian pada dokumen ini. Sistematika penulisan penelitian ini adalah sebagai berikut.

BAB 1 PENDAHULUAN

Bab ini membahas uraian mengenai latar belakang permasalahan, identifikasi masalah, menentukan maksud dan tujuan, batasan masalah, metodologi penelitian dan sistematika penulisan.

BAB 2 TINJAUAN PUSTAKA

Bab ini membahas berbagai konsep dasar dan teori-teori yang terkait dengan analisis sentimen berdasarkan aspek, teknik *preprocessing*, metode untuk pembobotan fitur, pengenalan *LVQ* sebagai metode klasifikasi, bahasa pemrograman, dan perangkat lunak lain yang digunakan.

BAB 3 ANALISIS DAN PERANCANGAN

Bab ini berisi tentang analisis dan perancangan sistem yaitu meliputi analisis masalah, analisis data masukan, analisis proses *preprocessing*, klasifikasi, analisis data keluaran, analisis kebutuhan fungsional dan *non* fungsional, perancangan antar muka, serta jaringan semantik.

BAB 4 IMPLEMENTASI DAN PENGUJIAN

Bab ini berisi tentang hasil dari keseluruhan tahap analisis dan perancangan yang meliputi implementasi data masukan, implementasi perangkat keras dan

perangkat lunak yang digunakan, implementasi antarmuka, serta pengujian dan hasil pengujian fungsionalitas sistem dan pengujian akurasi metode

BAB 5 KESIMPULAN DAN SARAN

Bab ini berisi hasil dari penelitian yang telah dilakukan sesuai dengan tujuan yang sudah ditetapkan, disertai dengan saran untuk pengembangan penelitian selanjutnya.