

# 1. Pendahuluan

## 1.1. Latar Belakang

Produk merupakan sesuatu yang ditawarkan oleh produsen terhadap konsumen. Setiap produk yang dikeluarkan oleh produsen, konsumen berhak memberikan opini atau pendapat mengenai suatu produk tersebut. Pendapat atau opini oleh konsumen tersebut dapat dijadikan tolak ukur bagi konsumen maupun produsen. Sebagai contoh, calon konsumen akan melihat terlebih dahulu opini-opini dari orang lain mengenai suatu produk sebelum membelinya. Berbeda dengan konsumen, produsen dapat menjadikan opini sebagai tolak ukur untuk evaluasi produk yang telah diproduksi [3]. Akan tetapi banyaknya opini yang ada pada suatu produk membuat konsumen maupun produsen kesulitan dalam menemukan opini mana yang termasuk opini positif dan opini mana yang termasuk opini negatif. Selain itu, konsumen maupun produsen kesulitan untuk menemukan fitur dari suatu produk mana yang paling banyak mendapat opini positif maupun negatif sebagai bahan evaluasi mereka. Maka dari itu diperlukan adanya sistem yang dapat menangani hal tersebut.

Pada penelitian tugas akhir kali ini, akan dibangun sistem yang dapat menentukan polaritas opini dari suatu fitur produk berdasarkan nilai kemiripan atau *similarity* suatu kata. Sistem yang akan dibangun menggunakan pendekatan Word2Vec untuk proses klasifikasi opini dimana Word2Vec merupakan model representasi kata dalam bentuk vektor, dan pada penelitian tugas akhir ini Word2Vec digunakan untuk melakukan klasifikasi opini pada fitur suatu produk [1]. Pendekatan Word2Vec dipilih pada penelitian tugas akhir ini sebab Word2Vec dapat mempelajari hubungan antar kata-kata secara otomatis dan dapat mempelajari berbagai macam *text corpus* sebagai data latih [4]. Word2Vec juga telah digunakan dalam beberapa penelitian terkait *sentiment analysis* seperti pada riset yang melakukan penelitian terkait *sentiment* pada Twitter [5], serta pada penelitian *sentiment movie review* berikut [6]. Penggunaan model Word2Vec pada penelitian tugas akhir ini juga diterapkan pada pengelompokan fitur produk atau *aspect grouping*, sehingga fitur produk tertentu dapat dikelompokkan dan dapat

diketahui jumlah opini *positive* maupun opini *negative* dari suatu fitur produk tersebut. Selain menggunakan pendekatan Word2Vec, penelitian tugas akhir ini juga menggunakan pendekatan WordNet pada klasifikasi opini. Hal ini dilakukan karena kedua pendekatan tersebut sama-sama memiliki nilai kemiripan atau *similarity* antar kata yang berbasis *graph* yang akan digunakan pada proses klasifikasi opini pada penelitian tugas akhir ini.

## **1.2. Perumusan Masalah**

Rumusan masalah dari penelitian tugas akhir ini adalah sebagai berikut:.

1. Bagaimana cara melakukan klasifikasi opini dari fitur produk berdasarkan nilai kemiripan atau *similarity* menggunakan Word2Vec dan WordNet?
2. Bagaimana cara mengelompokkan fitur produk serta mengetahui jumlah opini *positive* dan *negative* berdasarkan kelompok yang ditentukan menggunakan Word2Vec?

## **1.3. Tujuan**

Tujuan dari penelitian tugas akhir ini adalah sebagai berikut:

1. Menganalisis hasil penerapan Word2Vec dan WordNet pada klasifikasi opini produk.
2. Menganalisis pengelompokan fitur produk serta menganalisis jumlah opini *positive* dan *negative* berdasarkan kelompok yang ditentukan menggunakan Word2Vec.

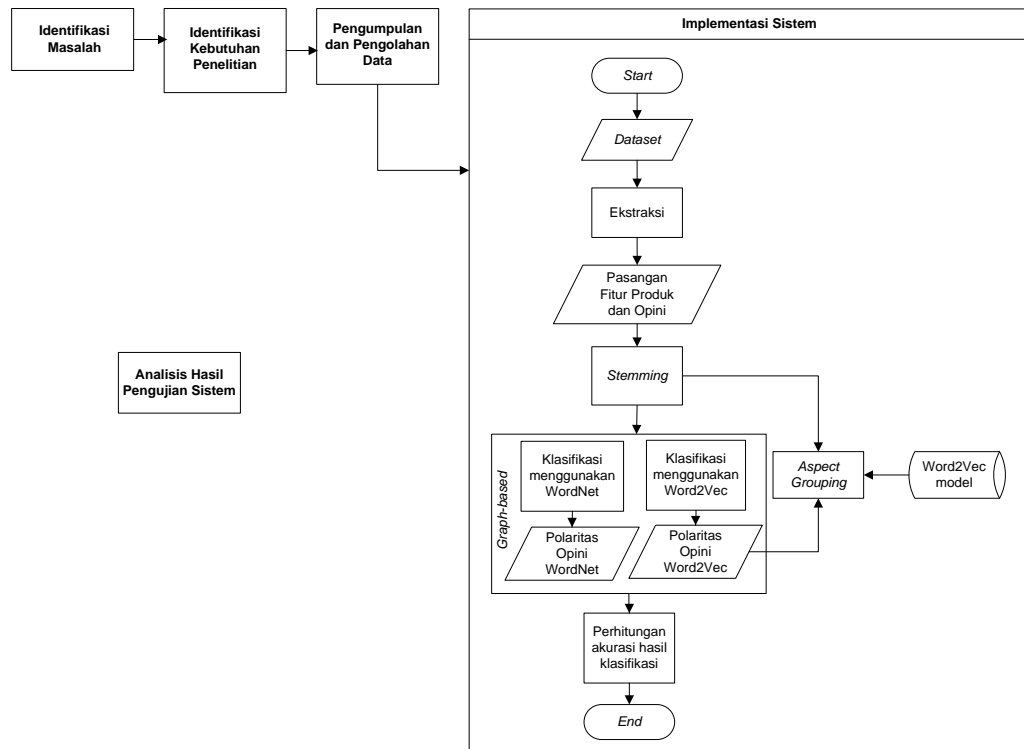
## **1.4. Batasan Masalah**

Adapun batasan masalah dari penelitian tugas akhir ini adalah sebagai berikut:

1. Proses ekstraksi opini dan fitur produk yang dilakukan pada penelitian tugas akhir ini menggunakan metode *double propagation*.
2. Penelitian tugas akhir ini tidak melakukan *preprocessing* terlebih dahulu.
3. *Dataset* yang digunakan merupakan *review* produk *camera* dalam bahasa Inggris dengan format *.txt* yang pernah digunakan pada penelitian dengan judul paper "*Mining and Summarizing Customers Reviews*" dari Minqing Hu dan Bing Liu.
4. Penelitian tugas akhir ini berfokus pada proses klasifikasi.

## 1.5. Metodologi Penyelesaian Masalah

Adapun metode penyelesaian masalah yang akan dilakukan pada penelitian tugas akhir ini ialah sebagai berikut:



Gambar 1-1 Metode Penelitian Tugas Akhir

### 1.5.1. Identifikasi Masalah

Terdapat beberapa masalah yang melatarbelakangi penelitian ini diantaranya ekstraksi kalimat *review* agar menghasilkan pasangan opini dan fitur produk. Selain itu, proses klasifikasi seperti penentuan polaritas dari suatu opini fitur produk juga teridentifikasi sebagai masalah dari penelitian ini, sebab nilai polaritas suatu opini akan memengaruhi akurasi dari proses klasifikasi yang akan dilakukan. Opini produk yang tidak tersusun berdasarkan fitur produk juga diidentifikasi sebagai masalah pada penelitian ini.

### 1.5.2. Identifikasi Kebutuhan

Kebutuhan yang akan digunakan pada penelitian tugas akhir ini dibagi menjadi dua kebutuhan, yaitu kebutuhan penelitian dan kebutuhan sistem. Adapun penjelasan dari masing-masing kebutuhan akan dipaparkan pada sub-bab berikut:

### 1.5.2.1. Identifikasi Kebutuhan Penelitian

Penelitian yang akan dilakukan membutuhkan cara penyelesaian masalah dengan menggunakan beberapa metode. Kebutuhan dalam penelitian tugas akhir ini disesuaikan dengan masalah yang ada, dimana masalah utama yang akan diselesaikan pada penelitian ini adalah bagaimana melakukan klasifikasi opini dari suatu fitur produk. Maka dari itu dibutuhkan metode-metode yang akan digunakan untuk menyelesaikan masalah-masalah tersebut. Pada penelitian tugas akhir ini, klasifikasi opini akan menggunakan nilai kemiripan atau *similarity* dari model Word2Vec.

### 1.5.2.2. Identifikasi Kebutuhan Sistem

Penelitian tugas akhir ini akan diimplementasikan dengan membangun sistem untuk menyelesaikan masalah. Sistem yang akan dibangun pada penelitian ini membutuhkan beberapa metode yang dapat membantu menyelesaikan masalah. Ekstraksi yang dilakukan pada sistem yang akan dibangun menerapkan metode *double propagation*, sedangkan untuk pembersihan kata opini dan fitur produk yang akan diklasifikasi menggunakan *stemming*.

### 1.5.3. Pengumpulan dan Pengolahan Data

Data yang digunakan pada penelitian tugas akhir ini merupakan *dataset* yang berupa *review* produk dalam bahasa Inggris dengan format *.txt* yang pernah digunakan pada penelitian dengan judul paper "*Mining and Summarizing Customers Reviews*" dari Minqing Hu dan Bing Liu.

Tabel 1-1 Dataset Produk Review [7]

No	Nama Produk	Jumlah Kalimat
1.	Canon Powershot SD500	229
2.	Canon G3	597
3.	Hitachi routers	312
4.	MicroMP3	1012
5.	Nikon coolpix 4300	346
6.	Nokia 6600	554

### 1.5.4. Implementasi Sistem

Kebutuhan-kebutuhan yang telah terpenuhi akan menjadi bahan untuk implementasi sistem. Implementasi sistem di sini dilakukan dengan melalui

beberapa tahap terlebih dahulu. Data yang sudah dikumpulkan berupa *dataset* akan digunakan pada tahap ekstraksi menggunakan *double propagation*. Proses ekstraksi akan menghasilkan pasangan opini dan fitur produk yang selanjutnya akan dilakukan proses *stemming* terlebih dahulu sebelum melalui proses klasifikasi. Setelah menerapkan *stemming* pada opini dan fitur produk, selanjutnya pasangan opini dan fitur produk tersebut masuk ke selanjutnya yaitu tahap klasifikasi. Tahap klasifikasi dilakukan dengan mencari nilai kemiripan atau *similarity* dari dua metode yakni dengan menggunakan model Word2Vec serta menggunakan perhitungan path distance dari WordNet. Hasil klasifikasi nantinya berupa fitur produk beserta polaritas opini dari kalimat *review*.

#### **1.5.5. Pengujian Sistem**

Pengujian sistem dilakukan dengan menggunakan sistem yang telah dibangun. Sistem yang akan diuji dibuat skenarionya terlebih dahulu. Skenario pengujian berisi rencana pengujian terhadap sistem untuk mengetahui tingkat keberhasilan setiap metode yang digunakan pada penelitian. Pengujian yang akan dilakukan nantinya akan menghasilkan hasil pengujian untuk dilakukan analisis.

#### **1.5.6. Analisis Hasil Pengujian Sistem**

Pengujian yang telah dilakukan sesuai dengan skenario akan menghasilkan hasil pengujian yang akan dianalisis. Analisis hasil pengujian sistem nantinya akan digunakan sebagai bahan evaluasi penelitian.

### **1.6. Sistematika Penulisan**

Sistematika penulisan pada Tugas Akhir ini adalah sebagai berikut.

#### **1. BAB 1 – Pendahuluan**

Pada Bab 1 menjelaskan tentang latar belakang penulisan Tugas Akhir ini kemudian menjelaskan perumusan masalah, tujuan serta batasan masalah yang ada pada pembuatan Tugas Akhir ini.

#### **2. BAB 2 – Tinjauan Pustaka**

Bab 2 menjelaskan tentang semua landasan teori yang digunakan dalam membangun Tugas Akhir.

#### **3. BAB 3 – Perancangan Sistem**

Pada Bab 3 menjelaskan seluruh rancangan sistem yang akan digunakan sebagai tolak ukur analisis untuk proses ekstraksi, klasifikasi dan *aspect grouping*.

4. BAB 4 – Pengujian dan Analisis

Pada Bab 4 berisi hasil analisis yang dilakukan berdasarkan rancangan yang telah dibuat untuk mengetahui perkembangan yang terjadi pada hasil evaluasi yang dilakukan dengan menggunakan perhitungan akurasi.

5. BAB 5 – Kesimpulan dan Saran

Bab 5 berisi kesimpulan yang didapat dari analisis yang telah dilakukan dan berisi saran untuk pengembangan sistem selanjutnya.