

# **BAB 1**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Virus *corona* adalah sebuah keluarga virus yang ditemukan pada manusia dan hewan. Sebagian virusnya dapat menginfeksi manusia serta menyebabkan berbagai penyakit, mulai dari penyakit umum seperti flu, hingga penyakit-penyakit yang lebih fatal, seperti *Middle East Respiratory Syndrome* (MERS) dan *Severe Acute Respiratory Syndrome* (SARS).

Pada saat ini seluruh dunia tengah mengalami serangan dari virus *corona* (COVID-19) termasuk di Indonesia. Di Indonesia virus *corona* (COVID-19) pertama kali dideteksi pada tanggal 20 maret 2020 hingga sekarang. Efek dari virus *corona* ini mengakibatkan semua aktifitas dibatasi untuk mencegah penyebaran virus *corona*, termasuk aktifitas pendidikan. Di Indonesia, sudah kurang lebih tiga bulan, semua aktifitas pendidikan, mulai dari tingkat dasar dan menengah hingga tingkat perguruan tinggi dilakukan secara daring, tanpa tatap muka antara pengajar dan peserta didik.

Pembelajaran daring, serta merta menyadarkan kita akan potensi luar biasa *internet* yang belum dimanfaatkan sepenuhnya dalam berbagai bidang, termasuk bidang pendidikan. Tanpa batas ruang dan waktu, kegiatan pendidikan bisa dilakukan kapanpun dan dimanapun. Terlebih lagi, di era dimana belum ada kepastian kapan pandemi ini akan berakhir, sehingga pembelajaran daring adalah kebutuhan mutlak yang harus dipenuhi oleh seluruh masyarakat Indonesia.

Namun, dibalik setiap sisi positif suatu hal, pastilah tersimpan sisi negatif, atau setidaknya kemungkinan buruk yang bisa saja terjadi. Meskipun secara formal kegiatan pendidikan masih bisa dilakukan secara daring, namun karena siswa dan mahasiswa harus belajar di rumah, pendidikan karakter selama masa pandemi ini, rasanya menjadi sedikit terabaikan. Khususnya seorang ibu harus memastikan anak-anak dan seluruh anggota keluarganya tetap berada di rumah dan membuat suasana nyaman. Termasuk juga menyediakan makanan bergizi dan seimbang. Ibu pula yang bertugas mengingatkan untuk memakai masker, menjaga jarak, dan selalu mencuci tangan. Ibu dengan anak usia sekolah juga mesti mendampingi anak-anak untuk belajar di rumah.

Berdasarkan permasalahan diatas, penelitian ini mencoba melakukan analisis sentimen dengan klasifikasi peran ibu rumah tangga dan ibu bekerja di masa pandemi COVID-19 di Pasir Pengaraian. Data tersebut akan diproses dengan *text minning* untuk menghindari kata yang kurang sempurna kemudian mengelompokkan data hasil reaksi kuesioner ke dalam dua klasifikasi yaitu positif dan negatif. Klasifikasi data ini menggunakan metode *K-Nearest Neighbor* (K-NN).

## **1.2 Rumusan Masalah**

1. Bagaimana hasil klasifikasi hasil reaksi tentang peran ibu rumah tangga dan ibu bekerja di masa pandemi COVID-19 dengan metode *K-Nearest Neighbor* (K-NN)?

2. Apakah metode *K-Nearest Neighbor* (K-NN) dapat digunakan dalam pengklasifikasian analisis sentimen peran ibu rumah tangga dan ibu bekerja dimasa pandemi COVID-19?

### **1.3 Batasan Masalah**

1. Data yang dianalisis adalah data dari hasil keusioner yang disebarakan melalui media sosial
2. Metode yang digunakan untuk klasifikasi adalah *K-Nearest Neighbor* (K-NN)
3. Jumlah data responden kuesioner yang digunakan adalah 100 responden.

### **1.4 Tujuan dan Manfaat Penelitian**

#### 1.4.1 Tujuan Penelitian

1. Membuktikan bahwa metode *K-Nearest Neighbor* (K-NN) dapat digunakan dalam pengklasifikasian *sentiment analysis* peran ibu rumah tangga dan ibu bekerja dimasa pandemi COVID-19
2. Dapat mengetahui tingkat efisiensi metode *K-Nearest Neighbor* (K-NN) dalam *sentiment analysis*

#### 1.4.2 Manfaat Penelitian

Adapun manfaat dari penelitian adalah sebagai berikut :

1. Memberikan informasi sentiment positif atau negatif tentang peran ibu rumah tangga dan ibu bekerja dimasa pandemi COVID-19

2. Memberikan pelajaran yang lebih banyak tentang *sentiment analysis*

## 1.5 Metodologi Penelitian

Tahapan-tahapan yang akan dilakukan pada pelaksanaan penelitian adalah sebagai berikut:

1. Studi Literatur

Pada tahap ini penulis mengumpulkan bahan referensi berkaitan dengan data minning, *text minning*, *sentiment analysis*, metode *K-Nearest Neighbour* (K-NN) dari berbagai jurnal, skripsi, buku, artikel dan berbagai sumber referensi lainnya.

2. Analisis Masalah

Pada tahap ini dilakukan analisis untuk setiap informasi yang telah di peroleh dari tahap sebelumnya agar mendapatkan pemahaman akan masalah dan metode yang akan digunakan untuk menyelesaikan *sentiment analysis*.

3. Perancangan Sistem

Pada tahap ini dilakukan perancangan sistem sesuai dengan hasil dari tahap sebelumnya.

4. Implementasi

Pada tahap ini hasil dari analisis dan perancangan sistem akan di implementasikan ke dalam kode program.

## 5. Pengujian

Pada tahap ini dilakukan pengujian terhadap *sentiment analysis* para ibu rumah tangga dan ibu bekerja.

## 6. Dokumentasi dan Penyusunan Laporan

Pada tahap terakhir membuat dokumentasi dan menyusun laporan hasil dari analisis dan implementasi dari penelitian tersebut.

### 1.6 Sistematika Penulisan

Sistematika penulisan dari skripsi ini terdiri dari lima bagian utama sebagai berikut :

#### **Bab 1 PENDAHULUAN**

Bab Ini Berisi Latar Belakang, Rumusan Masalah Tujuan Penelitian, Batasan Masalah, Manfaat Penelitian, Metodologi Penelitian Dan Sistematika Penulisan.

#### **Bab 2 LANDASAN TEORI**

Bab Ini Berisi Teori-Teori Yang Digunakan Pada Penelitian Ini. Teori-Teori Yang Berhubungan Dengan *sentiment analysis*, data *minning*, *text minning*, Metode *K-Nearest Neighbour* (K-NN).

#### **Bab 3 METODE PENELITIAN**

Bab Ini Berisi Tahapan–Tahapan Dalam Pengumpulan Data, Perancangan Sistem Perumusan Masalah Dan Analisa.

#### **Bab 4 ANALISA DAN PERANCANGAN**

Bab Ini Berisi Analisa Dan Perancangan Aplikasi Penerapan Metode *K-Nearest Neighbour* (K-NN) untuk menganalisis tentang ibu rumah tangga dan ibu bekerja dimasa pandemi COVID-19.

#### **Bab 5 IMPLEMENTASI DAN PENGUJIAN**

Bab ini berisi implementasi dari analisa dan perancangan dan pengujian pada aplikasi yang berhasil dibangun.

#### **Bab 6 PENUTUP**

Bab ini berisi rangkuman dari hasil penelitian yang telah dilakukan dan saran – saran untuk pengembangan aplikasi atau penelitian selanjutnya.

## **BAB 2**

### **LANDASAN TEORI**

#### **2.1 *Sentiment Analysis***

Analisis sentimen atau *opinion mining* merupakan proses memahami, mengekstrak dan mengolah data tekstual secara *otomatis* untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat opini. Analisis sentimen dilakukan untuk melihat pendapat atau kecenderungan opini terhadap sebuah masalah atau obyek oleh seseorang, apakah cenderung berpandangan atau beropini negatif atau positif. Salah satu contoh penggunaan analisis sentimen dalam dunia nyata adalah identifikasi kecenderungan pasar dan opini pasar terhadap suatu obyek barang. Besarnya pengaruh dan manfaat dari analisis sentimen menyebabkan penelitian dan aplikasi berbasis analisis sentimen berkembang pesat [1].

Menurut Medhat et al (2014, 1093) Analisis sentimen adalah suatu bidang yang sedang berlangsung dalam penelitian berbasis teks. Analisis sentimen atau opini *mining* adalah kajian tentang cara untuk memecahkan masalah dari opini masyarakat, sikap dan emosi suatu entitas, dimana entitas tersebut dapat mewakili individu, peristiwa atau topik [2].

#### **2.2 *Text Mining***

*Text Mining* adalah proses menambang data yang berupa teks dimana sumber data biasanya didapatkan dari dokumen dan tujuannya adalah mencari kata-kata yang dapat mewakili isi dari dokumen sehingga dapat dilakukan analisis keterhubungan antar dokumen tersebut. *Text mining* mengekstrak informasi

berguna dari sumber data melalui identifikasi dan eksplorasi yang tidak dalam bentuk *database record*, melainkan dalam data teks yang tidak terstruktur. *Preprocessing* adalah tahap proses awal *text mining* terhadap teks untuk mempersiapkan teks menjadi data yang dapat diolah lebih lanjut. Sekumpulan karakter yang bersambungan (teks) harus dipecah-pecah menjadi unsur yang lebih berarti. Suatu dokumen dapat dipecah menjadi bab, sub-bab, paragraf, kalimat, kata dan bahkan suku kata[3].

*Text mining* juga dikenal sebagai data *mining* teks atau penemuan pengetahuan dari *database* tekstual. Sesuai dengan buku *The Text Mining Handbook* [10], *text mining* dapat didefinisikan sebagai suatu proses menggali informasi dimana seorang *user* berinteraksi dengan sekumpulan dokumen menggunakan *tools* analisis yang merupakan komponen-komponen dalam data *mining*. Tujuan dari *text mining* adalah untuk mendapatkan informasi yang berguna dari sekumpulan dokumen. Jadi, sumber data yang digunakan dalam *text mining* adalah sekumpulan teks yang memiliki format yang tidak terstruktur atau minimal semi terstruktur. Adapun tugas khusus dari *text mining* antara lain yaitu pengkategorisasian teks dan pengelompokkan teks [4].

### **2.3 Data Mining**

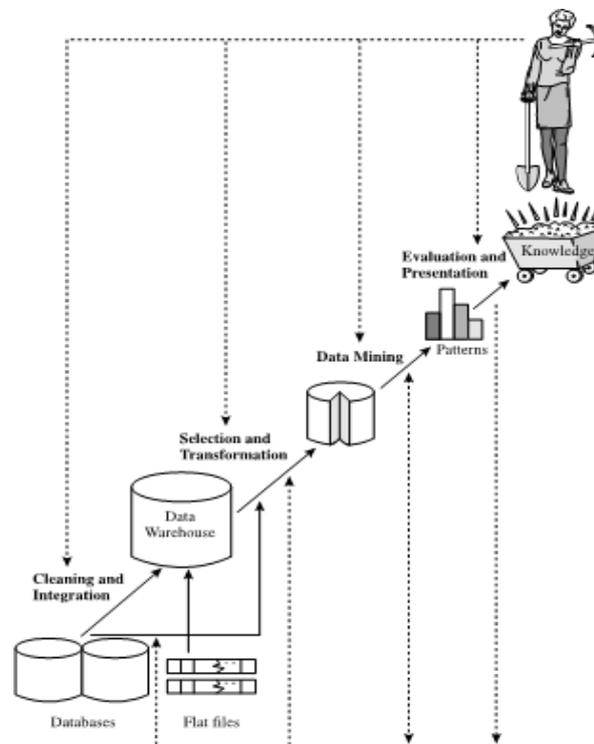
Data *mining* merupakan proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakut dari berbagai *database* besar, keluaran dari data *mining* bisa dipakai untuk memperbaiki pengambilan keputusan dimasa depan.

Berdasarkan definisi-definisi yang telah disampaikan, hal penting yang terkait dengan Data *Mining* adalah:

1. Data *mining* merupakan suatu proses otomatis terhadap data yang sudah ada.
2. Data yang akan diproses berupa data yang sangat besar.
3. Tujuan data *mining* adalah mendapatkan hubungan atau pola yang akan mungkin memberikan indikasi yang bermanfaat [5].

### 2.3.1 Tahap-Tahap Data *Mining*

Sebagai suatu rangkaian proses, data *mining* dapat dibagi menjadi beberapa tahap proses yang diilustrasikan pada Gambar 2.1. Tahap-tahap tersebut bersifat interaktif, pemakai terlibat langsung atau dengan perantaraan *knowledge Base* [6].



Tahap-tahap data *minning* adalah sebagai berikut:

a. Pembersihan data (*data cleaning*)

Pembersihan data merupakan proses menghilangkan-kan *noise* dan data yang tidak konsisten atau data tidak relevan.

b. Integrasi data (*data integration*)

Integrasi data merupakan penggabungan data dari berbagai *database* ke dalam satu *database* baru.

c. Seleksi data (*data selection*)

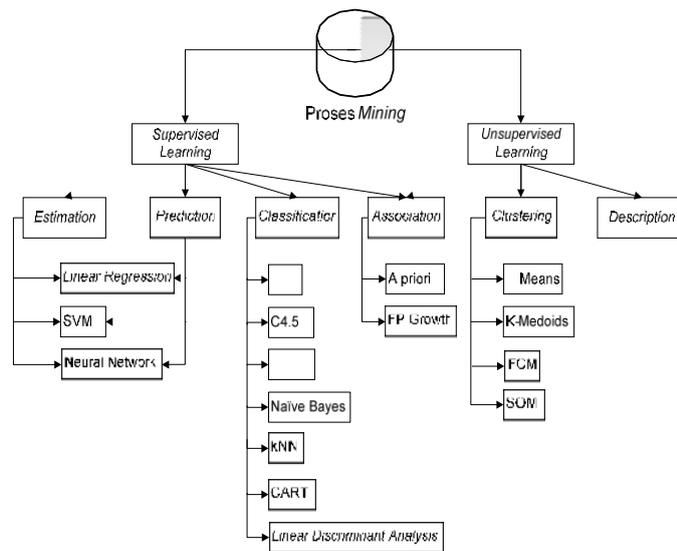
Data yang ada pada *database* sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari *database*.

d. Transformasi data (*data transformation*)

Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam data *minning*.

e. Proses *minning*

Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data. Beberapa metode yang dapat digunakan berdasarkan pengelompokan data *minning* dapat dilihat pada Gambar 2.2.



**Gambar 2.2 Beberapa Metode Data Mining**

f. Evaluasi pola (*pattern evaluation*)

Untuk mengidentifikasi pola-pola menarik ke dalam *knowledge based* yang ditemukan.

g. Presentasi pengetahuan (*knowledge presentation*)

Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna.

**2.4 Corona Virus (COVID-19)**

*Corona* virus merupakan virus RNA *strain* tunggal positif, berkapsul dan tidak bersegmen. *Corona* virus tergolong *ordo Nidovirales*, keluarga *Coronaviridae*. Struktur *corona* virus membentuk struktur seperti kubus dengan protein dan S berlokasi di permukaan virus. Protein S atau *spike* protein merupakan salah satu protein antige utama virus dan merupakan struktur utama untuk penulisan gen. Protein S ini berperan dalam penempelan dan masuknya virus kedalam sel *host* (interaksi protein S dengan reseptornya di sel inang).

*Corona* virus bersifat sensitif terhadap panas dan secara efektif dapat di inaktifkan oleh desinfektan mengandung klorin, pelarut lipid dengan suhu 56°C selama 30 menit, eter, alkohol, asam perioksiasetat, detergen non-ionik, formalin, oxidizing agent dan kloform. Klorheksidin tidak efektif dalam menonaktifkan virus [7].

## 2.5 Metode *K-Nearest Neighbor* (K-NN)

*K-Nearest Neighbor* (K-NN) adalah suatu metode yang menggunakan algoritma *super-vised* dimana hasil dari *query instance* yang baru diklasifikasikan berdasarkan mayoritas dari kategori pada K-NN. Tujuan dari algoritma ini adalah mengklasifikasikan obyek baru berdasarkan atribut dan *training sample*. *Classifier* tidak menggunakan model apapun untuk dicocokkan dan hanya berdasarkan pada memori. Diberikan titik *query* akan ditemukan sejumlah *k* obyek atau (titik *training*) yang paling dekat dengan titik *query*. Klasifikasi menggunakan *voting* terbanyak diantara klasifikasi dari obyek. Algoritma K-NN menggunakan klasifikasi ketetanggaan sebagai nilai prediksi dari *query instance* yang baru.

Algoritma metode K-NN sangatlah sederhana, bekerja berdasarkan jarak terpendek dari *query instance* ke *training sample* untuk menentukan K-NN-nya. *Training sample* diproyeksikan ke ruang berdimensi banyak, dimana masing-masing dimensi merepresentasikan fitur dari data. Ruang ini dibagi menjadi bagian-bagian berdasarkan klasifikasi *training sample*. Sebuah titik pada ruang ini ditandai kelas *c* jika kelas *c* merupakan klasifikasi yang paling banyak ditemui pada *k* buah tetangga terdekat dari titik tersebut. Dekat atau jauhnya tetangga

biasanya dihitung berdasarkan *Euclidean Distance* yang direpresentasikan pada persamaan 1 sebagai berikut [8] :

$$\sqrt{\sum_{k=1}^d (a_k - b_k)^2} \quad (1)$$

Dimana :

$D(a,b)$  = jarak sekular dari kedua *vector a* dan *b* dari matriks dengan ukuran dimensi

## 2.6 XAMPP

*XAMPP* adalah perangkat lunak bebas, yang mendukung banyak sistem operasi, merupakan kompilasi dari beberapa program. Fungsinya adalah sebagai *server* yang berdiri sendiri (*localhost*), yang terdiri atas program *Apache*, *HTTP Server*, *Mysql database*, dan penerjemah bahasa yang ditulis dengan bahasa pemrograman *PHP* dan *Perl* [9].

## 2.7 PHP

(*PHP Hypertext Preprocessor*) adalah bahasa pemrograman yang berjalan dalam sebuah *web server* dan berfungsi sebagai pengolah data pada sebuah *server*. *PHP* adalah bahasa pemrograman yang bersifat *open source* yang dapat digunakan di berbagai mesin ( *Linux*, *Unix*, *Macintosh*, *Windows*) dan dapat dijalankan secara runtime melalui console serta juga dapat menjalankan perintah-perintah sistem. *PHP* bisa terintegrasi dengan *database* seperti *Mysql* ataupun *Microsoft Acces* [10].

## 2.8 UML

UML (*Unified Modeling Language*) adalah salah satu alat bantu yang sangat handal di dunia pengembangan sistem yang berorientasi obyek. Hal ini disebabkan karena UML menyediakan bahasa pemodelan visual yang memungkinkan bagi pengembang sistem untuk membuat cetak biru atas visi mereka dalam bentuk baku, mudah dimengerti serta dilengkapi dengan mekanisme yang efektif untuk berbagi (*Sharing*) dan mengkomunikasikan rancangan mereka dengan yang lain.

Komponen-komponen yang diperlukan sebagai berikut:

### 1) *Use Case Diagram*

*Use case* adalah deskripsi fungsi dari sebuah sistem dari perspektif pengguna. *Use case* bekerja dengan cara mendeskripsikan tipikal interaksi antara *user* (pengguna) sebuah system dengan sistemnya sendiri melalui sebuah cerita bagaimana sebuah system dipakai dengan urutan langkah-langkah yang menerangkan antara pengguna disebut dengan aktor dan sistem mendeskripsikan urutan kejadian.

### 2) *Class Diagram*

*Class diagram* adalah diagram statis dari suatu aplikasi tidak hanya digunakan untuk menggambarkan dan mendokumentasikan berbagai aspek sistem tetapi untuk membangun kode eksekusi dari aplikasi perangkat lunak sebagai suatu set obyek yang memiliki atribut dan perilaku yang sama, kelas kadang disebut kelas obyek *Class* memiliki tiga area pokok yaitu :

- 1) Nama, kelas harus mempunyai sebuah nama.
- 2) Atribut, adalah kelengkapan yang melekat pada kelas. Nilai dari suatu kelas hanya bisa diproses sebatas atribut yang dimiliki.
- 3) Operasi, adalah proses yang dapat dilakukan oleh sebuah kelas, baik pada kelas itu sendiri ataupun kepada kelas lainnya.

### 3) *Activity Diagram*

*Diagram activity* menunjukkan aktivitas sistem dalam bentuk kumpulan aksi-aksi, bagaimana masing-masing aksi tersebut dimulai, keputusan yang mungkin terjadi hingga berakhirnya aksi. *Activity diagram* juga dapat menggambarkan proses lebih dari satu aksi dalam waktu bersamaan. “*Diagram activity* adalah aktifitas-aktifitas, obyek, state, transisi *state* dan *event*. Dengan kata lain kegiatan diagram alur kerja menggambarkan perilaku sistem untuk aktivitas.

### 4) *Sequence Diagram*

*Sequence diagram* adalah gambaran tahap demi tahap, termasuk kronologi (urutan) perubahan secara logis yang seharusnya dilakukan untuk menghasilkan sesuatu sesuai dengan *use case diagram* [12].

## 2.9 Penelitian Terdahulu

Penelitian terdahulu ini menjadi salah satu acuan penulis dalam melakukan penelitian sehingga penulis dapat memperkaya teori yang digunakan dalam mengkaji penelitian yang dilakukan. Dari penelitian terdahulu, penulis tidak menemukan penelitian dengan judul yang sama seperti judul penelitian penulis. Namun penulis mengangkat beberapa penelitian sebagai referensi dalam memperkaya bahan kajian pada penelitian penulis. Berikut merupakan penelitian

terdahulu berupa beberapa jurnal terkait dengan penelitian yang dilakukan penulis.

**Tabel. 2.1 Penelitian Terdahulu**

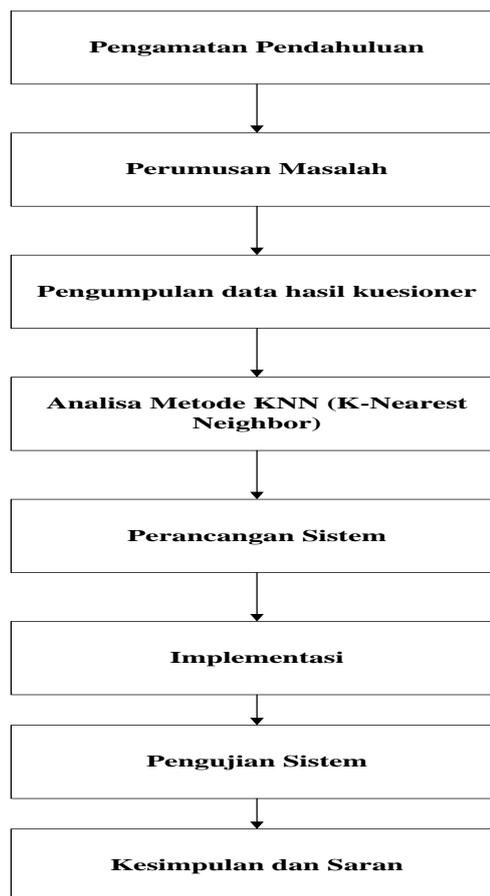
Nama Peneliti	Judul Penelitian	Hasil Penelitian
Dyarsa Singgih Pamungkas, Noor Ageng Setiyanto, Erlin Dolphine	Analisi Sentiment Pada Sosial Media Twiiter Menggunakan Naïve Bayes Classifier Terhadap Kata Kunci “Kurikulum 2013”	1. Klasifikasi tweet bersentimen lebih akurat jika data latih yang di gunakan semakin banyak dalam data pengetahuan. 2. Akurasi Naive Bayes Classifier memberikan hasil sebesar 91 % untuk 1000 data latih yang diberikan. 3. Fungsi N-gram kata dapat meningkatkan analisis sentimen. 4. Jika hasil Vmap setiap kategori sama akan menghasilkan kategori tidak tersentimentkan. 5. Jika hasil Vmap pada tweet ada yang berjumlah nol(0) maka data pengetahuan kurang. 6. Analisis tidak berjalan maksimal terhadap bahasa asing dan bahasa daerah.
Juen Ling, I Putu Eka N.Kencana, Tjokorda Bagus Okta	Analisis Sentimen Menggunakan Metode <i>Naïve Bayes Classifier</i> Dengan Seleksi Fitur <i>Chi Square</i>	Berdasarkan hasil yang diperoleh dapat disimpulkan bahwa kemunculan frekuensi fitur pada kategori yang diharapkan dan kategori yang tidak diharapkan memiliki peranan penting dalam seleksi fitur

		<p><i>Chi Square</i>, oleh karena itu seleksi fitur <i>Chi Square</i> baik digunakan dalam penyeleksian fitur dibandingkan dengan metode <i>frequency-based</i>. Serta pembangunan sistem analisis sentimen menggunakan metode NBC dengan bahasa pemrograman Java memperoleh akurasi sebesar 83% dan rata-rata harmonik sebesar 90,713%. Terdapat kesalahan klasifikasi karena pada data uji terdapat fitur yang muncul pada bukan kategorinya. Untuk penelitian selanjutnya yaitu dapat dibandingkan hasil seleksi fitur dari <i>chi square</i> terhadap hasil seleksi fitur dari <i>Mutual Information</i> berdasarkan segi waktu komputasi dan segi ketepatan klasifikasi. Serta menggabungkan teknik pembelajaran mesin NBC dengan beberapa model <i>n</i>-gram untuk meneliti hasil ketepatan klasifikasi yang diperoleh.</p>
--	--	--

### **BAB 3**

## **METODOLOGI PENELITIAN**

Tahapan penelitian yang akan dilakukan dan penyelesaian masalah terhadap *sentiment analysis* para ibu rumah tangga dan ibu bekerja dimasa pandemi COVID-19 dengan metode *K-Nearest Neighbor* (K-NN). Adapun tahapan metodologi yang dilakukan selama penelitian dapat dilihat pada Gambar 3.1, yang mana merupakan proses yang dimulai dari studi literatur hingga diperoleh kesimpulan.



**Gambar 3.1 Tahapan Metodologi Penelitian**

Pembuatan tugas akhir ini terbagi menjadi beberapa tahap pengerjaan yang tertera sebagai berikut:

### **3.1 Pengamatan Pendahuluan**

Pengamatan pendahuluan merupakan tahapan awal yang dilakukan dalam penelitian ini, yang menggunakan metode *K-Nearest Neighbor* (K-NN) yang dijadikan sebagai penelitian studi pustaka dalam penelitian Tugas Akhir ini. Pada penelitian metode ini yaitu menggunakan data *sentiment analysis* para ibu rumah tangga dan ibu bekerja di masa pandemi COVID-19 menggunakan metode *K-Nearest Neighbor* (K-NN)".

### **3.2 Perumusan Masalah**

Berdasarkan hasil dari tahapan pengamatan pendahuluan sebelumnya, maka tahapan selanjutnya adalah tahapan perumusan masalah. Pada tahapan perumusan masalah akan dirumuskan masalah yang dianggap sebagai penelitian dalam Tugas Akhir ini. Permasalahan-permasalahan yang menjadi rumusan masalah dalam penelitian ini didapatkan dari penelitian terkait data pengamatan pendahuluan sebelumnya serta hasil dari keusioner. Solusi yang didapatkan pada tahapan perumusan masalah ini yang akan menjadi judul penelitian Tugas Akhir ini "*Sentiment Analysis* para Ibu Rumah Tangga dan Ibu dan Ibu Bekerja di Masa Pandemi COVID-19 Menggunakan Metode *K-Nearest Neighbor* (K-NN)".

### **3.3 Pengumpulan Data**

Pengumpulan data adalah tahapan-tahapan yang bertujuan dalam memperoleh data-data informasi yang berhubungan dengan penelitian Tugas Akhir ini. Pada tahapan pengumpulan data ini juga berguna untuk mengumpulkan

semua kebutuhan data yang akan diproses nantinya menggunakan metode “*K-Nearest Neighbor* (K-NN)”. Dalam pengumpulan data ini data yang dikutip adalah sebagai berikut:

#### 1. Studi Literatur

Dalam proses penelitian, diperlukan pengumpulan pengetahuan dengan cara mempelajari literatur dari beberapa bidang ilmu yang berhubungan dengan data *minning*, *text minning*, *sentiment analysis*, metode *K-Nearest Neighbour* (K-NN) dari berbagai jurnal, skripsi, buku, artikel dan berbagai sumber referensi yaitu diantaranya:

- a. Pengumpulan informasi mengenai *sentiment analysis* ibu rumah tangga dan ibu beketja
- b. Pengumpulan informasi terkait metode *K-Nearest Neighbor* (K-NN)
- c. Pengumpulan data dari jurnal dan buku-buku.
- d. Mengumpulkan hasil keusioner dari 100 responden.

#### 2. Analisis Kebutuhan Sistem

Dalam analisa sistem bertujuan mengidentifikasi sistem yang akan dirancang, yang meliputi perangkat lunak serta perangkat keras. Tahapan-tahapan yang menyusun analisa sistem ini adalah analisis data yang dipakai, spesifikasi kebutuhan sistem, spesifikasi pengguna, perancangan basis data, dan perancangan antarmuka. Adapun spesifikasi dari perangkat keras (*hardware*) dan perangkat lunak software (*software*) yang digunakan sebagai berikut :

a. Perangkat keras (*hardware*), antara lain :

Prosesor : Intel(R) Core(TM) i3-4200U CPU *up to*

2.48 Ghz *Memory* (RAM) : 4.00 GB

*System type* : 64-bit *Operating system*

Harddisk : 500 GB

b. Perangkat Lunak (*software*), antara lain :

Sistem Operasi : *Windows 7*

*Tool* : Xampp, Notepad ++, *Google Chrome*

### **3.4 Pengujian**

Pengujian merupakan sebuah tahapan yang memperlihatkan apakah prediksi tingkat akurasi dari penelitian sesuai dengan yang diinginkan atau tidak.

### **3.5 Kesimpulan dan Saran**

Tahapan terakhir adalah menarik kesimpulan dari hasil penelitian yang didapatkan dalam sentimen analisis para ibu rumah tangga dan ibu bekerja di masa pandemi COVID-19. Pada tahapan ini juga berisikan saran peneliti bagi pembaca untuk melakukan pengembangan terhadap penelitian ini kedepannya.