

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Keluhan pasien merupakan tahap awal pemeriksaan yang dilakukan oleh dokter untuk mengumpulkan informasi terkait kondisi yang sedang dirasakan oleh pasien agar dokter dapat menyimpulkan hasil diagnosis dari pasien tersebut. Dalam rekam medis elektronik, data pengobatan pasien dicatat ke dalam catatan klinis naratif (Xu *et al.*, 1996). Rekam medis elektronik merupakan sebuah sistem informasi kesehatan yang berisi data keluhan pasien termasuk data sosial dan data medis pasien (Mada, 2017).

Keluhan pasien adalah salah satu jenis data klinis yang penting dalam catatan medis. Informasi yang terdapat dalam data keluhan pasien saat ini belum dikelola dengan baik secara digital. Data keluhan pasien sering dicatat dalam catatan klinis sebagai teks bebas tanpa adanya proses lanjutan seperti ekstraksi informasi untuk mengambil informasi yang penting dalam data tersebut. Tanpa adanya proses ekstraksi informasi pada data keluhan pasien dapat menghambat proses pengambilan informasi terhadap data-data inti pada keluhan pasien jika sewaktu-waktu informasi tersebut diperlukan.

Ekstraksi informasi merupakan sebuah proses yang dilakukan untuk mendapatkan informasi inti atau utama dari data-data yang tidak terstruktur menjadi terstruktur (Ismaya, 2014). Proses ekstraksi yang dilakukan secara manual membutuhkan waktu yang lama karena dilakukan dengan cara membaca teks dan mengidentifikasi data yang relevan, kemudian disimpan ke dalam basis data.

Berdasarkan permasalahan di atas, dibutuhkan sebuah sistem yang mampu mengekstrak informasi yang ada pada data keluhan pasien dengan memetakan teks pada keluhan tersebut. Keluhan pasien tersebut diklasifikasikan dalam sebelas *slot/field*, yaitu keluhan utama, onset, keluhan lain, keterangan, frekuensi serangan, sifat serangan, durasi, lokasi, perjalanan penyakit, riwayat pengobatan sebelumnya, dan akibat gangguan yang timbul (Ratnasari *et al.*, 2014).

Dalam penelitian ini, penulis menggambarkan sistem ekstraksi informasi klinis dengan pendekatan statistik yang dapat mengekstrak keluhan-keluhan pada pasien dari narasi klinis agar nantinya dapat diolah oleh Sistem Pendukung Keputusan Klinis (SPKK). Metode yang

biasa digunakan dalam pendekatan statistik ini adalah *Support Vector Machine* (SVM), *Hidden Markov Model* (HMM), dan *Conditional Random Fields* (CRF). Berdasarkan hasil kajian jurnal terkait perbandingan kekurangan dan kelebihan dari tiga metode tersebut seperti rangkuman analisis pada Tabel 1.1, maka didapatkan metode yang digunakan dalam penelitian ini adalah metode *Support Vector Machine* (SVM).

Tabel 1.1 Perbandingan Metode

No	Metode	Peneliti	Kelebihan	Kekurangan
1.	<i>Support Vector Machine</i> (SVM)	(Pratama & Bambang, 2015) (Nugroho <i>et al.</i> , 2003) (Quora, 2015)	<ul style="list-style-type: none"> <li>- Paling cocok untuk mengklasifikasi suatu <i>pattern</i></li> <li>- Klasifikasi token yang dapat membagi teks menjadi token-token</li> <li>- Setiap token dapat menentukan bagian dari <i>slot</i> (<i>classifier</i>)</li> <li>- <i>State of the art</i> untuk permasalahan klasifikasi dan <i>named entity recognition</i></li> <li>- Klasifikasi teks dengan akurasi paling tinggi sebesar 87% berdasarkan penelitian oleh Aqsath</li> </ul>	<ul style="list-style-type: none"> <li>- Sulit digunakan dalam masalah yang berskala besar (jumlah sampel yang diolah)</li> <li>- Tidak dapat mempertimbangkan dependensi <i>state-to-state</i> dibandingkan dengan CRF</li> </ul>
2.	<i>Hidden Markov Model</i> (HMM)	(Gozali, Maharani & A, 2010) (Seymore <i>et al.</i> , 2014) (Quora, 2014)	<ul style="list-style-type: none"> <li>- <i>Key Topic Sentence Extracting</i> dan <i>Key Word Indexing</i></li> <li>- Fondasi statistiknya kuat</li> <li>- Sangat cocok untuk domain pada bahasa alami</li> <li>- Penanganan untuk data yang baru lebih efisien</li> </ul>	<ul style="list-style-type: none"> <li>- Hanya bergantung pada label pada kata sebelumnya</li> <li>- Tidak dapat menangani data dengan baik dibandingkan dengan CRF</li> </ul>
3.	<i>Conditional Random Fields</i> (CRF)	(Praisar <i>et al.</i> , 2016) (Quora, 2015)	<ul style="list-style-type: none"> <li>- Pelabelan data terstruktur atau berurutan</li> <li>- Dapat mempertimbangkan dependensi <i>state-to-</i></li> </ul>	<ul style="list-style-type: none"> <li>- Membutuhkan waktu lebih lama daripada HMM</li> </ul>

			state dibandingkan dengan SVM	
--	--	--	-------------------------------	--

Pada Tabel 1.1 di atas merupakan penjelasan perbandingan mengenai kelebihan dan kekurangan dari metode *Support Vector Machine* (SVM), *Hidden Markov Model* (HMM), dan *Conditional Random Fields* (CRF).

## 1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, masalah yang dapat dirumuskan adalah bagaimana membangun sistem yang dapat mengekstrak informasi dari bahasa alami keluhan pasien agar lebih terstruktur dengan menggunakan metode statistik *Support Vector Machine* (SVM) ?

## 1.3 Batasan Masalah

Adapun batasan masalah dari penelitian ini adalah sebagai berikut:

- a. Data keluhan pasien diperoleh dari dokter.
- b. Data yang akan digunakan adalah keluhan pasien Riwayat Penyakit Sekarang (RPS), bukan Riwayat Penyakit Dahulu (RPD) atau Riwayat Penyakit Keluarga (RPK).

## 1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah menghasilkan sistem yang dapat mengekstrak informasi yang ada pada data teks narasi keluhan pasien dengan memetakan teks pada keluhan tersebut ke dalam objek-objek Riwayat Penyakit Sekarang (RPS) agar nantinya hasil dari penelitian ini dapat diintegrasikan dengan Sistem Pendukung Keputusan Klinis (SPKK).

## 1.5 Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah sebagai berikut :

- a. Data keluhan pasien menjadi lebih terstruktur sehingga nantinya dapat diintegrasikan dengan Sistem Pendukung Keputusan Klinis (SPKK).
- b. Penulisan keluhan pasien sesuai dengan bahasa alami pasien karena pada kenyataan sekarang ini keluhan pasien sesuai dengan bahasa dokter.
- c. Data keluhan pasien terdokumentasi dengan baik.

## 1.6 Metodologi Penelitian

Adapun langkah-langkah yang diterapkan untuk membuat sistem dalam penelitian ini adalah sebagai berikut:

### a. Tahap Pengambilan Data

#### 1. Wawancara

Wawancara dilakukan dengan dr. Rahma, dr. Ari Erna, dan dr. Olan, selaku dokter di Klinik Unisia Polifarma UII. Wawancara ini dilakukan untuk mendapatkan informasi mengenai keluhan pasien.

#### 2. Observasi

Tahap observasi dalam penelitian ini adalah mengumpulkan informasi data keluhan pasien di Apotek Unisia Polifarma UII. Data yang dikumpulkan adalah bahasa alami dari hasil rekaman suara pasien yang diperoleh saat wawancara dengan dokter.

### b. Tahap Pembuatan Sistem

#### 1. Analisis Kebutuhan

Tahap ini bertujuan untuk mengetahui kebutuhan-kebutuhan yang diperlukan untuk membuat sistem yang dapat mengekstrak informasi pada keluhan pasien. Kebutuhan tersebut berupa data-data keluhan dari pasien dan dokter. Dalam penelitian ini, dilakukan diskusi dengan dokter dan dosen pembimbing.

#### 2. Perancangan

Perancangan merupakan penggambaran perencanaan sistem yang akan dibuat agar lebih terstruktur dan jelas. Rancangan yang dibuat berupa desain tampilan antarmuka sistem.

#### 3. Implementasi

Implementasi merupakan tahapan yang dilakukan untuk menerapkan atau membangun sistem yang telah dirancang sebelumnya.

#### 4. Pengujian

Tahap pengujian ini merupakan tahapan terakhir, tahapan ini dilakukan untuk mengecek kesalahan yang terdapat pada sistem yang telah selesai dibuat.

## 1.7 Sistematika Penulisan

Sistematika penulisan laporan berguna untuk memudahkan dalam memahami laporan tugas akhir ini. Dalam penyusunan tugas akhir ini, sistematika penulisan terdiri dari lima bab, yaitu :

### **BAB I Pendahuluan**

Berisi penjelasan mengenai latar belakang masalah, rumusan masalah, batasan masalah, tujuan, manfaat, metodologi penelitian, dan sistematika penulisan mengenai penelitian yang dilakukan.

### **BAB II Landasan Teori**

Berisi uraian mengenai teori-teori pendukung terkait dengan topik penelitian yang dilakukan.

### **BAB III Metodologi Penelitian**

Berisi langkah-langkah penyelesaian dan analisis pada penelitian yang dilakukan.

### **BAB IV Hasil dan Pembahasan**

Berisi hasil dari tampilan sistem dan pembahasan sistem secara keseluruhan.

### **BAB V Kesimpulan dan Saran**

Berisi kesimpulan yang menyimpulkan apakah hasil penelitian yang dilakukan telah sesuai dengan tujuan penelitian atau belum serta berisi saran yang mendukung agar penelitian dapat dikembangkan lagi oleh peneliti lain yang ingin memperbaiki kekurangan dari penelitian ini.