

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Dewasa ini, di era teknologi begitu banyak hal mengenai sistem informasi konvensional yang sudah berubah menjadi sistem informasi digital. Sistem informasi yang sudah digital akan menghasilkan pencarian informasi yang lebih efektif dan efisien. Data yang akan diteliti pada penelitian ini adalah surat keputusan. Surat keputusan adalah surat yang dikeluarkan oleh instansi atau organisasi yang diwakili oleh pimpinan yang tertinggi yang berisi pernyataan memutuskan sesuatu hal yang berhubungan dengan tertib organisasi yang bersangkutan. Namun, surat keputusan yang dimiliki sebuah perusahaan cukup banyak. Pencarian surat keputusan pada ratusan bahkan ribuan surat keputusan akan memakan waktu yang sangat lama. Maka, dibutuhkan sistem pencarian arsip surat keputusan agar dalam pencarian arsip bisa lebih mudah dan cepat. Selain itu, dengan mengarsipkan informasi surat fisik kedalam bentuk digital akan membuat surat keputusan lebih terjaga kemanan surat keputusan dari kerusakan kertas yang akan terjadi seiring bertambahnya waktu.

Proses ekstraksi informasi (*Information Extraction* atau IE) adalah sebuah proses untuk mendapatkan informasi inti dari dokumen tidak terstruktur menjadi dokumen terstruktur. Adapun data yang terstruktur adalah data yang telah terorganisir sehingga mudah dalam suatu pencarian data sedangkan tidak terstruktur adalah data yang belum terorganisi [1]. Pada penelitian sebelumnya proses ekstraksi informasi pada surat masuk yang menggunakan metode *naïve bayes* oleh Chandra Ratiwi menghasilkan akurasi yang sangat tinggi yakni 96,96%. Dokumen yang digunakan pada penelitian tersebut berupa surat. Penelitian tersebut bermaksud untuk mengekstraksi informasi dengan pengklasifikasian setiap kalimat yang terdapat pada surat masuk, karena surat masuk tidak selalu identic dengan format yang sudah ada [2]. Sedangkan, pada penelitian ekstraksi informasi pada dokumen teks novel oleh Arbi Insani menggunakan metode rule based menghasilkan hasil

akurasi yang juga sangat baik sebesar 100%. Dokumen yang digunakan pada penelitian tersebut berupa novel yang mana informasi yang akan diambil berupa data judul, penulis, tahun terbit, penerbit, ISBN, dan E-ISBN yang mana data yang digunakan memiliki format yang sama [3]. Pada penelitian *Long Short Term Memory* (LSTM) tentang identifikasi klikbait sebelumnya menghasilkan akurasi 82% [4]. Pada penelitian LSTM tentang prakira cuaca menghasilkan 91,8%[5][5]. Pada penelitian pengenalan dialek menghasilkan akurasi saat diuji dengan Bahasa sunda dialek garut sebesar 100%, diuji dengan Bahasa sunda yang dialeknnya bukan garut sebesar 77,778%, dan untuk orang yang bukan sunda sebesar 66,667% [6]. Pada penelitian tentang pengenalan dialek manado menghasilkan akurasi sebesar 87% [7].

Namun, dikarenakan surat keputusan yang digunakan dengan format yang sebagian sama dan sebagian lagi berbeda maka, dipenelitian ini menggunakan salah satu metode *Deep Learning LSTM* yang mana dengan metode tersebut akan mengolah proses yang lebih baik tanpa harus data masukan tersebut berformat sama atau tetap serta dikarenakan dokumen yang digunakan berformat berbeda, penelitian ini menggunakan metode LSTM, pada penelitian sebelumnya tentang ekstraksi informasi yang tidak terstruktur dengan dokumen makalah ilmiah yang diakses *online* [8], maka pada penelitian ini ingin mengetahui kinerja serta akurasi dari metode LSTM.

1.2 Perumusan Masalah

Berdasarkan latar belakang yang telah dikemukakan, diambil rumusan masalah yaitu bagaimana mendapatkan identitas informasi dari sebuah surat keputusan yang memiliki format yang berbeda menggunakan metode *Long Short Term Memory*.

1.3 Maksud dan Tujuan

Penelitian ini bermaksud untuk menerapkan metode ekstraksi informasi pada surat keputusan. Sedangkan tujuan yang ingin dicapai dari penelitian ini adalah mengetahui akurasi dari metode *Long Short Term Memory*.

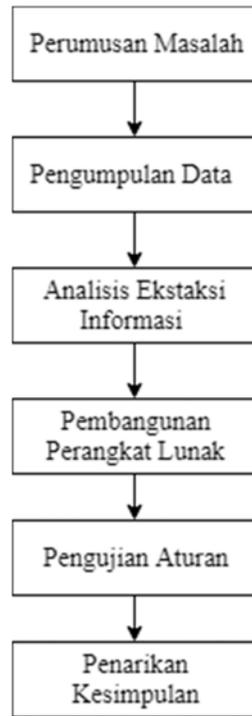
1.4 Batasan Masalah

Agar masalah yang sedang ditinjau lebih terarah dan mencapai tujuan yang telah ditentukan, maka akan dibatasi masalah sebagai berikut :

1. Dokumen yang digunakan merupakan surat keputusan berformat *.jpg yang di konversi menggunakan OCR menjadi *.txt.
2. Surat keputusan hanya surat yang berbahasa Indonesia.
3. Data Masukan yang akan diolah pada sistem berupa dokumen berformat *.jpg.
4. Keluaran yang dihasilkan sistem adalah nomor induk pegawai, nama pegawai, nomor surat, dan tanggal penetapan surat keputusan.
5. Program penelitian yang dibuat menggunakan *library tesseract* dan *library keras*.

1.5 Metode Penelitian

Metode penelitian yang dipakai adalah metode Deskriptif. Metode Deskriptif merupakan yang mempelajari masalah-masalah yang ada serta tata cara kerja yang berlaku. Pada metode ini menghimpun data, menyusunnya secara sistematis, faktual dan cermat. Metode penelitian ini memiliki dua tahapan yaitu pengumpulan data dan pembangunan perangkat lunak. Alur Penelitian yang digunakan dapat dilihat pada Gambar 1.1 Alur Penelitian.



Gambar 1. 1 Alur Penelitian

1.5.1 Perumusan Masalah

Tahap pertama dalam penelitian adalah perumusan masalah. Peneliti merumuskan masalah ataupun kendala pada proses ekstraksi informasi pada surat keputusan.

1.5.2 Implementasi dan Pengujian

Implementasi dan Pengujian adalah pengumpulan data. Metode pengumpulan data yang dilakukan pada penelitian ini yaitu :

a. Kajian Pustaka

Kajian pustaka dilakukan dengan membaca jurnal, paper, buku dan situs online terkait dengan topik ekstraksi informasi, metode *Long Short Term Memory* sistem dan yang berkaitan dengan topik penelitian.

b. Pengumpulan Data

Pada tahap ini diambil pengumpulan surat keputusan dengan 50 sampel surat keputusan. Data sampel ini digunakan sebagai data training saat melakukan ekstraksi informasi pada surat keputusan.

1.5.3 Analisis Ekstraksi Informasi

Adapun analisis yang dilakukan pada surat keputusan adalah sebagai berikut.

1. Identifikasi Surat keputusan

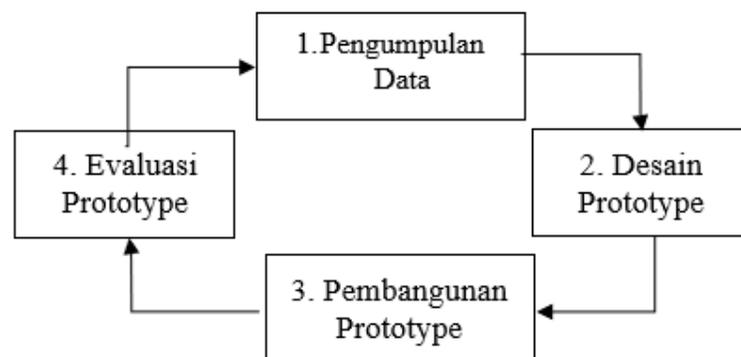
Pada tahap analisis identifikasi surat keputusan adalah mengamati data surat keputusan yang digunakan pada penelitian dan menentukan data surat keputusan yang akan diambil.

2. Analilis Metode

Pada tahap ini menganalisis metode yang akan digunakan.

1.5.4 Implementasi Pembangunan Perangkat Lunak

Metode analisis data dalam pembuatan perangkat lunak menggunakan model prototyping karena model ini melakukan pendekatan secara sistematis dan berurutan dalam pembangunan perangkat lunak dapat dilihat pada Gambar 1.2 Diagram Model Prototyping.



Gambar 1. 2 Diagram Model Prototyping [9]

Langkah-langkah pada model prototyping adalah sebagai berikut [9] :

1. Pengumpulan kebutuhan

Pada proses pengumpulan kebutuhan, pada penelitian ini data dikumpulkan dari instansi dan internet.

2. Proses desain

Desain berfokus pada representasi dari aspek perangkat lunak dari sudut pandang pengguna; ini mencakup input, proses output. Pada penelitian ini dibuat 3 bagian program yang mana terdiri dari bagian pertama *pre-processing*, bagian kedua *learning*, bagian ketiga hasil dari data yang akan diambil dari ekstraksi informasi.

3. Membangun prototype

Langkah selanjutnya adalah pembuatan atau perancangan produk yang sesungguhnya.

4. Evaluasi dan perbaikan

Pada tahap ini dilakukan evaluasi pada penelitian yang sudah dilakukan.

1.5.5 Pengujian Sistem

Tahap selanjutnya adalah pengujian sistem yang telah dibangun. Pengujian ini dilakukan dengan menggunakan pengujian *black box* dan pengujian hasil ekstraksi informasi menggunakan pengecekan manual.

1.5.6 Pernarikan Kesimpulan

Tahapan akhir ini adalah menarik kesimpulan terhadap aplikasi yang telah dibangun.

1.6 Sistematika Penulisan

Sistematika penulisan ini disusun untuk memberikan gambaran umum tentang penelitian yang akan dijalankan. Sistematika penulisan tugas akhir ini adalah sebagai berikut.

BAB 1 PENDAHULUAN

Bab ini merupakan bab yang membahas latar belakang masalah, perumusan masalah, maksud dan tujuan, batasan masalah, metodologi penelitian, serta sistematika penulisan tentang ekstraksi informasi.

BAB 2 LANDASAN TEORI

Bab ini membahas berbagai konsep dasar dan teori-teori yang berkaitan dengan surat keputusan, ekstraksi informasi, kecerdasan buatan, bahasa pemrograman python, kajian pendekatan aturan *long short term memory*, uraian singkat UML, flowchart, jaringan semantic, basis data, dan pengujian.

BAB 3 ANALISIS DAN PERANCANGAN

Bab ini menjelaskan analisis dan perancangan sistem yang dibangun. Analisis sistem terdiri dari analisis masalah, analisis data, analisis aturan, sistem yang sedang berjalan, analisis kebutuhan non fungsional yang terdiri dari analisis kebutuhan perangkat keras, analisis kebutuhan perangkat lunak.

BAB 4 IMPLEMENTASI DAN PENGUJIAN PERANGKAT LUNAK

Bab ini menjelaskan implementasi dari perangkat lunak yang dibangun. Implementasi perangkat lunak dilakukan berdasarkan kebutuhan analisis dan perancangan perangkat lunak yang sudah dilakukan. Dari hasil implementasi kemudian dilakukan pengujian berdasarkan pendekatan aturan *long short term memory* pada analisis kebutuhan perangkat lunak yang menjelaskan apakah sudah benar-benar sesuai dengan analisis dan perancangan yang telah dilakukan.

BAB 5 KESIMPULAN DAN SARAN

Bab ini berisi tentang kesimpulan hasil penelitian yang telah dilakukan dan saran yang diharapkan dapat dikembangkan untuk penelitian selanjutnya.

