

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Penelitian merupakan suatu kegiatan penyelidikan dengan aktif, tekun dan secara sistematis bertujuan untuk menemukan, menginterpretasikan dan membantah fakta-fakta yang ada. Kegiatan ini menghasilkan suatu pengetahuan yang lebih mendalam mengenai suatu peristiwa, tingkah laku, teori dan hukum [1]. Selain itu, penelitian juga dapat menghasilkan suatu pengetahuan baru yang sebelumnya belum pernah ada dan dapat dipertanggung jawabkan.

Laporan hasil penelitian tersebut dapat dituliskan secara sistematis dalam sebuah artikel ilmiah. Maksud dari laporan yang sistematis adalah laporan yang disusun dengan mengikuti struktur dan format yang berlaku dalam suatu jurnal ilmiah. Jurnal ilmiah adalah sebuah publikasi yang diterbitkan secara berkala oleh suatu organisasi profesi atau institusi akademik yang memuat artikel-artikel yang merupakan produk pemikiran ilmiah secara empiris (artikel hasil penelitian) maupun secara logis (artikel hasil pemikiran) dalam bidang ilmu tertentu [2].

Ringkasan merupakan penyajian secara singkat dari suatu dokumen dengan mengambil beberapa kalimat penting yang mewakili isi dokumen dan mempertahankan sudut pandang pengarang. Tujuan dibuatnya suatu ringkasan adalah untuk membantu memahami dan mengetahui isi dari dokumen[3]. Untuk mendapatkan ringkasan dapat dilakukan secara manual, yaitu dengan membaca keseluruhan isi dokumen. Akan tetapi waktu yang dibutuhkan pun mungkin cukup

lama. Bisa saja dokumen tersebut memiliki kumpulan teks dengan jumlah yang banyak.

Peningkatan teknologi informasi telah memberi terobosan baru dalam peringkasan suatu dokumen. Lama waktu yang sebelumnya menjadi kendala dalam mendapatkan ringkasan dapat diatasi dengan lebih cepat. Aplikasi peringkasan otomatis dapat digunakan untuk mendapatkan ringkasan dokumen dengan cepat tanpa harus membaca seluruh isi dokumen dengan menggunakan teknik *text mining* [4][5].

Penelitian mengenai peringkasan teks otomatis telah dilakukan sejak tahun 1958 oleh Luhn dan masih terus berkembang hingga saat ini [6][5]. Peringkasan teks otomatis merupakan sebuah alat untuk mendapatkan bentuk yang lebih singkat dari suatu dokumen teks dengan menggunakan aplikasi yang dijalankan pada komputer [7]. Tujuannya adalah untuk mendapatkan informasi dengan mengambil sebagian besar hal yang dianggap penting dan ditampilkan dalam bentuk yang ringkas tanpa membaca keseluruhan dokumen.

Beberapa penelitian telah membahas peringkasan teks otomatis dengan menggunakan berbagai metode yang telah diterapkan. Salah satu metode untuk meringkas suatu dokumen adalah *Non-Negative Matrix Factorization*. Metode ini dikategorikan ke dalam metode *unsupervised* yang tidak membutuhkan data latihan sebagai proses pembelajaran untuk meringkas serta dalam membuat suatu ringkasan dengan membangkitkan fitur semantik dari suatu dokumen.

Penelitian menggunakan metode *Non-negative Matrix Factorization* telah diterapkan pada peringkasan teks otomatis, diantaranya : *Peringkasan Dokumen dengan*

Metode Non-Negative Matrix Factorization [4], *Document Summarization Using NMF And Pseudo Relevance Feedback Based On K-Means Clustering* [8] dan *Automatic Summarization of The Arabic Documents using NMF : A Preliminary Study* [5].

Pada penelitian ini akan mengembangkan kembali penelitian sebelumnya dengan dokumen yang menjadi objek penelitian adalah jurnal ilmiah berbahasa Indonesia. Serta mengidentifikasi faktor yang mempengaruhi mengapa *term* yang terdapat pada hasil ringkasan menjadi lebih penting dibandingkan dengan yang lain. Maka akan dibangun aplikasi peringkas teks otomatis jurnal berbahasa Indonesia dengan judul “**Implementasi Metode *Non-Negative Matrix Factorization* Untuk Peringkas Teks Otomatis Jurnal Bahasa Indonesia**”.

1.2 Perumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, maka dapat dirumuskan menjadi beberapa rumusan masalah yang diantaranya :

1. Bagaimana mengimplementasikan metode *non-negative matrix factorization* dalam meringkas jurnal bahasa Indonesia ?
2. Bagaimana kinerja metode *non-negative matrix factorization* dalam meringkas jurnal bahasa Indonesia ?
3. Bagaimana mengidentifikasi faktor yang mempengaruhi suatu *term* yang terdapat pada ringkasan menjadi lebih penting dibandingkan dengan yang lain ?

1.3 Tujuan Penelitian

Adapun tujuan penelitian ini diantaranya :

1. Mengimplementasikan metode *non-negative matrix factorization* dalam meringkas jurnal bahasa Indonesia.
2. Mengetahui kinerja metode *non-negative matrix factorization* dalam meringkas jurnal bahasa Indonesia.
3. Mengidentifikasi faktor yang mempengaruhi *term* yang terdapat pada ringkasan menjadi lebih penting dibandingkan dengan yang lain.

1.4 Batasan Masalah

Agar penelitian yang dilakukan lebih terfokus dan tidak menyimpang dari tujuan yang ingin dicapai, maka perlu beberapa batasan masalah yang diantaranya :

1. Jurnal yang digunakan memiliki format pdf.
2. Jurnal ilmiah berbahasa Indonesia.
3. Jumlah jurnal yang diringkas hanya satu jurnal (*single document*).
4. Aplikasi dirancang menggunakan UML (*Unified Modeling Language*).
5. Menggunakan metode *non-negative matrix factorization*.
6. Ukuran file yang digunakan maksimal 2mb.
7. Kriteria jurnal yang digunakan, yaitu *header* dan *footer* dihilangkan, bab dan sub bab dihilangkan, identitas jurnal dihilangkan, judul jurnal dihilangkan.
8. Ringkasan yang dihasilkan berupa beberapa kalimat asli yang diambil dari dokumen tersebut.

9. Aplikasi yang akan dibangun berbasis web menggunakan bahasa pemrograman Python.

1.5 Metodologi Penelitian

1.5.1 Metode Pengumpulan Data

Metode yang digunakan dalam pengumpulan data pada penelitian ini terdiri dari 2 tahap yaitu :

1. Observasi

Teknik pengumpulan data dengan mengadakan penelitian dan peninjauan langsung terhadap objek penelitian.

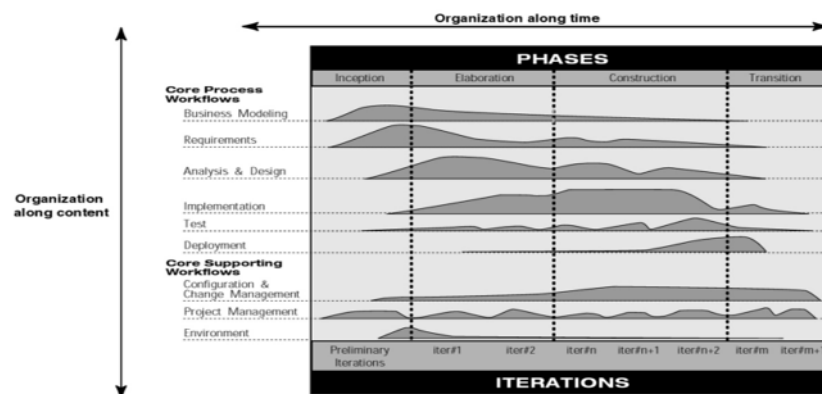
2. Studi Literatur

Pengumpulan data dengan cara mengumpulkan literatur, jurnal dan bacaan-bacaan yang berkaitan dengan penelitian.

1.5.2 Metode Pengembangan Perangkat Lunak

Metode pengembangan perangkat lunak yang akan digunakan untuk membangun aplikasi dalam penelitian ini menggunakan metode RUP (*Rational Unified Process*). RUP (*Rational Unified Process*) adalah pendekatan pengembangan perangkat lunak yang dilakukan berulang-ulang/iteratif, fokus pada arsitektur sistem, lebih diarahkan berdasarkan penggunaan kasus [9].

Untuk proses iterasi dari RUP yang digambarkan dalam grafik dapat dilihat pada Gambar 1.1 :



Gambar 1. 1 Grafik iterasi RUP [10].

Sedangkan untuk proses pengulangan iteratif dari metode RUP ini dapat dilihat pada Gambar 1.2 :



Gambar 1. 2 RUP Model [9].

RUP memiliki empat tahap yang dapat dilakukan secara iteratif juga. Berikut penjelasan mengenai tahapan-tahapan yang terdapat pada metode RUP [9][10]:

1. *Inception* (Permulaan)

Tahap ini lebih pada pemodelan proses bisnis yang dibutuhkan (*business modeling*) dan mendefinisikan

kebutuhan akan sistem yang akan dibuat (*requirements*).

Berikut ini adalah tahap yang dibutuhkan pada tahap ini :

- Gambaran dokumen : Gambaran umum kebutuhan utama proyek, Fitur kunci dan Masalah utama.
- Inisialisasi *Use case* (minimal 10%-20%).
- Inisialisasi kasus bisnis, yang mana termasuk konteks bisnis, kriteria sukses dan perkiraan biaya.
- Inisialisasi perkiraan resiko.
- Rencana proyek, menampilkan fase dan iterasi.
- Model bisnis, jika dibutuhkan.
- Satu atau beberapa *prototype*.

Jika pada tahap akhir ini target yang diinginkan tidak dicapai maka dapat dibatalkan atau diulang kembali setelah dirancang ulang agar kriteria yang diinginkan dapat dicapai. Batas/tonggak objektif digunakan untuk mendeteksi apakah sebuah kebutuhan akan sistem dapat diimplementasikan atau tidak.

2. *Elaboration* (Perluasan/Perencanaan)

Tahap ini memfokuskan diri pada perencanaan arsitektur sistem. Tahap ini juga dapat mendeteksi apakah arsitektur sistem yang diinginkan dapat dibuat atau tidak. Mendeteksi resiko yang mungkin terjadi. Pada tahap *elaboration* ini lebih pada analisis dan desain sistem serta implementasi sistem

yang fokus pada *prototype*. Berikut ini adalah tahapan yang dilakukan pada tahap *elaboration* :

- *Use case* setidaknya rampung dalam 80%, semua *use case* dan aktor telah diidentifikasi dan seluruh deskripsi *use case* sudah dibuat.
- Penambahan kebutuhan yang diantaranya kebutuhan non-fungsional dan beberapa kebutuhan yang tidak terkait dengan *use case* tertentu.
- Deskripsi arsitektur aplikasi.
- Sebuah *prototype* arsitektur yang bisa dijalankan.
- Daftar resiko dan kasus bisnis yang telah diperbaiki.
- Rencana pembangunan untuk keseluruhan aplikasi, termasuk rencana proyek yang masih kasar, menampilkan iterasi-iterasi dan kriteria evaluasi dari setiap iterasi.
- Sebuah kasus pengembangan terbaru untuk menentukan proses yang digunakan.
- Panduan penggunaan manual (Opsional).

Jika pada akhir tahap ini target yang diinginkan tidak dicapai maka dapat dibatalkan atau diulang kembali. Batas/tonggak arsitektur digunakan untuk mendeteksi apakah sebuah kebutuhan akan sistem dapat diimplementasikan atau tidak melalui pembuatan arsitektur.

3. *Construction* (Konstruksi)

Tahap ini fokus pada pengembangan komponen dan fitur-fitur sistem. Tahap ini lebih pada implementasi dan pengujian sistem yang fokus pada implementasi perangkat lunak pada kode program. Dalam tahap ini juga menghasilkan produk perangkat lunak dimana menjadi syarat dari *Initial Operational Capability Milestone* atau batas/tonggak kemampuan operasional awal. Hasil dari tahap konstruksi ini merupakan sebuah produk yang siap digunakan oleh pengguna, setidaknya terdiri dari :

- Aplikasi yang bisa terintegrasi dengan *platforms* yang telah ditentukan.
- *User manuals*.
- Deskripsi rilis.

4. *Transition* (Transisi)

Tahap ini lebih pada *deployment* atau instalasi sistem agar dapat dimengerti oleh pengguna. Tahap ini menghasilkan produk perangkat lunak dimana menjadi syarat dari *Initial Operational Capability Milestone* atau batas/tonggak kemampuan operasional awal. Aktifitas pada tahap ini termasuk pada pelatihan pengguna, pemeliharaan dan pengujian sistem apakah sudah memenuhi harapan pengguna. Adapun tujuan utama dari tahap transisi adalah sebagai berikut :

- Mencapai persetujuan dengan *stakeholder* bahwa penyebaran *baseline* sudah lengkap dan konsisten dengan evaluasi kriteria visi.
- Mencapai produk akhir dengan cepat dan biaya praktik yang efektif.

Produk perangkat lunak juga disesuaikan dengan kebutuhan yang didefinisikan pada tahap *inception*. Jika semua kriteria objektif terpenuhi maka dianggap sudah memenuhi *Product Release Milestone* (batas/tonggak peluncuran produk) dan pengembangan perangkat lunak selesai dilakukan.

Akhir dari keempat fase ini adalah produk perangkat lunak yang sudah lengkap. Keempat fase pada RUP dijalankan secara berurutan dan iteratif dimana setiap iterasi dapat digunakan untuk memperbaiki iterasi berikutnya [9].

1.6 Sistematika Penulisan

Penulisan skripsi ini dibagi kedalam lima bab, yang disusun berdasarkan sistematika sebagai berikut :

BAB I PENDAHULUAN

Dalam bab ini dijelaskan tentang Latar belakang masalah, Perumusan masalah, Tujuan penelitian, Batasan masalah, Metodologi penelitian dan Sistematika penulisan.

BAB II STUDI PUSTAKA

Pada bab ini berisi tentang penjelasan teori-teori yang menunjang untuk tugas akhir serta menyelesaikan permasalahan yang akan dikaji.

BAB III PERANCANGAN SISTEM

Pada bab ini akan membahas analisis dan perancangan untuk aplikasi yang akan dibangun.

BAB IV IMPLEMENTASI DAN PENGUJIAN

Bab ini membahas dan menguji perangkat lunak yang telah dibangun. Proses implementasi meliputi kebutuhan perangkat keras, kebutuhan perangkat lunak dan tampilan aplikasi yang dibangun. Kemudian, dilakukan tahap pengujian aplikasi meliputi identifikasi penerapan metode, *software* dan pembahasan hasil pengujian.

BAB V PENUTUP

Dalam bab ini akan dijelaskan mengenai beberapa kesimpulan dan saran-saran.

