

BAB I

PENDAHULUAN

1.1. Latar Belakang

Seiring perkembangan teknologi dan informasi yang pesat dimana semua orang dengan mudah mendapatkan informasi. Perkembangan teknologi dan informasi membawa dampak yang sangat signifikan di dalam kehidupan manusia dan di berbagai bidang, salah satunya dibidang pendidikan (Mahdavi dkk., 2014), dimana memungkinkan peserta didik maupun pengajar dapat memperoleh informasi tidak hanya bersumber dari buku-buku yang telah diterbitkan, tetapi dapat memperoleh informasi dari *e-book*, artikel ataupun jurnal baik yang berskala nasional hingga internasional yang diakses melalui internet. Namun adanya kemudahan tersebut memicu semakin besarnya peluang kemungkinan peserta didik ataupun pengajar melakukan tindakan-tindakan yang menyimpang contohnya plagiarisme.

Plagiarisme juga dapat dikategorikan berupa tulisan atau potongan tulisan maupun secara keseluruhan tidak dikutip dengan baik dan benar, baik tindakan tersebut sengaja ataupun tidak disengaja. Kasus plagiarisme ini sedang maraknya terjadi di Indonesia beberapa tahun terakhir ini (Soelistryo, 2011). Tak hanya pelajar, mahasiswa di Indonesia saja, tetapi bahkan dinegara maju pun tidak terlepas dari plagiarisme (Chao dkk.,2009) untuk menyelesaikan tugas harian, karya tulis ilmiah maupun tugas akhir. Hal ini dibuktikan dengan adanya beberapa penelitian seperti yang dilakukan oleh Mulyana (2010) menyatakan bahwa tindak plagiarisme yang dilakukan mahasiswa dalam penulisan skripsi adalah duplikasi judul, duplikasi isi skripsi, duplikasi teori , duplikasi data, dan duplikasi referensi. Penelitian lainnya yang dilakukan oleh Bhalerao dan Wilhelm (2015) menyatakan bahwa, sekitar 70% mahasiswa mengaku telah melakukan plagiarisme. Berdasarkan hasil pengumpulan data dengan metode *accidental sampling* di Fakultas Sains dan Teknologi Universitas Katolik Musi Charitas yang

melibatkan 62 responden, terdapat sekitar 80.6% mahasiswa, tahu dan mengerti tentang plagiarisme, 58.1% mahasiswa pernah mengutip kutipan orang lain tanpa mencantumkan sumbernya, 45.2% mahasiswa pernah melakukan plagiarisme untuk menyelesaikan tugas kuliahnya, 21% mahasiswa pernah melakukan plagiarisme untuk menyelesaikan ujian, dan 3.2% mahasiswa pernah melakukan plagiarisme untuk menyelesaikan tugas akhir atau skripsi.

Adapun berbagai cara yang digunakan untuk mengurangi tindakan plagiarisme salah satunya dengan membandingkan dokumen yang diajukan dengan dokumen lainnya. Jika tingkat kemiripan antar dokumen atau bagian dari dokumen tinggi, maka dapat diduga telah terjadi tindakan plagiarisme pada dokumen tersebut (Isa dan Abidin., 2013). Beberapa penelitian – penelitian yang dilakukan selama 10 tahun terakhir, ditemukan berbagai algoritma yang dapat digunakan untuk mengukur kemiripan dokumen, 2 diantaranya adalah algoritma *Vector Space Model*, dan algoritma *Rabin Karp*. Dengan melakukan analisis perbandingan performansi dari algoritma *vector space model* dan algoritma *rabin karp*, maka dapat diketahui cara kerja dan performansi dari kedua algoritma tersebut.

1.2. Perumusan Masalah

Berdasarkan latar belakang yang telah diuraikan diatas, permasalahan yang menjadi fokus dalam penelitian ini adalah membangun aplikasi yang dapat mengukur kemiripan dokumen menggunakan algoritma *Vector Space Model* dan algoritma *Rabin Karp* yang berbasis *web*.

1.3. Batasan Masalah

Mengingat adanya keterbatasan dalam hal biaya untuk melakukan penelitian dan untuk menghindari kompleksitas yang mungkin timbul dan memungkinkan akan memakan waktu yang lama dalam melakukan penelitian, maka diberikan batasan - batasan dalam melakukan penelitian sebagai berikut.

1. Domain masalah dalam penelitian ini dibatasi hanya untuk mengukur kemiripan dokumen proposal skripsi yang ada di Fakultas Sains dan Teknologi Universitas Katolik Musi Charitas, khususnya program studi Informatika,

sehingga tidak adanya verifikasi terhadap dokumen-dokumen eksternal yang ada di internet.

2. Dokumen yang diproses oleh aplikasi ini adalah dokumen teks dengan format *.docx* dan menggunakan bahasa Indonesia, sedangkan dokumen dengan format *.doc* tidak bisa digunakan dikarenakan isi dokumen tersebut tidak bisa dibaca, hal ini dibuktikan dengan beberapa penelitian yang dijadikan studi literatur yang tidak ada menggunakan format dokumen *.doc* di dalam penelitian.

1.4. Tujuan Dan Manfaat Penelitian

Adapun tujuan dan manfaat yang ingin dicapai dalam melakukan penelitian ini dapat dilihat sebagai berikut:

1.4.1 Tujuan Penelitian

Berdasarkan latar belakang dan permasalahan yang telah dibahas, maka penelitian ini bertujuan untuk menganalisis cara kerja antara algoritma *vector space model* dengan algoritma *rabin karp*, dalam mengukur kemiripan dokumen.

1.4.2 Manfaat Penelitian

Berdasarkan permasalahan dan tujuan dari penelitian ini, adapun manfaat yang diharapkan dari hasil penelitian ini yakni: menemukan algoritma yang lebih efisien antara algoritma *vector space model* dengan algoritma *rabin karp*, dalam mengukur kemiripan dokumen.

1.5. Metodologi Penelitian

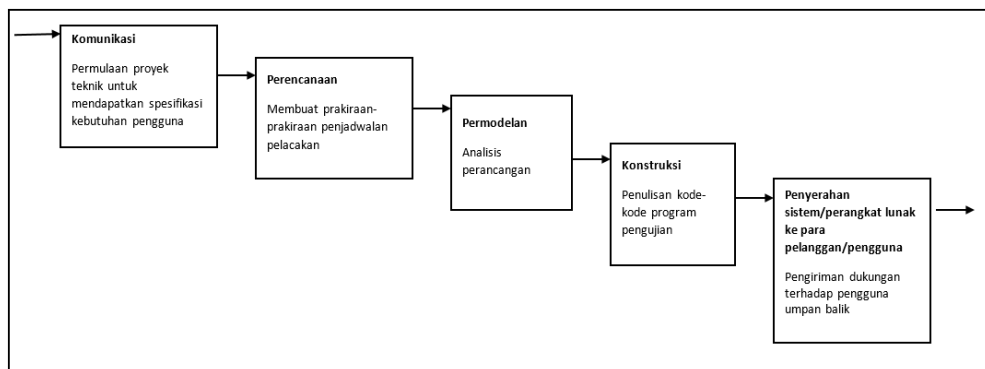
Metodologi penelitian yang digunakan dalam penelitian ini dapat dilihat sebagai berikut:

1.5.1 Jenis Penelitian

Berdasarkan klasifikasi penelitian berdasarkan tujuan dan manfaatnya, penelitian ini merupakan jenis penelitian terapan. Penelitian ini dilakukan untuk menentukan algoritma yang lebih efisien dalam mengukur tingkat kemiripan dokumen antara algoritma *vector space model* dan algoritma *rabin karp*.

1.5.2 Metode Pengembangan Sistem

Metode pengembangan sistem yang digunakan dalam penelitian ini adalah model *waterfall*. Model ini melakukan pendekatan kepada perkembangan perangkat lunak yang sistematis dan sekuensial yang berurutan untuk pengembangan perangkat lunak yang dimulai dari spesifikasi dan persyaratan pelanggan (*communication*) dan berlangsung melalui perencanaan (*planning*), pemodelan (*modeling*), konstruksi (*construction*), serta penyerahan sistem atau perangkat lunak ke para pelanggan atau pengguna (*deployment*), yang diakhiri dengan dukungan berkelanjutan pada perangkat lunak lengkap yang dihasilkan (Pressman, 2012). Adapun ilustrasinya dapat dilihat pada gambar 1.1.



Gambar 1.1 Model Waterfall
(Sumber : Pressman, 2012)

Adapun penjelasan mengenai fase-fase tersebut menurut Pressman (2012)

1. Komunikasi

Langkah ini dimaksudkan untuk memahami tujuan pelanggan atau pihak yang berkepentingan atas proyek perangkat lunak yang sedang dikembangkan dan mengumpulkan kebutuhan-kebutuhan yang akan membantu. Tahap ini dilaksanakan dengan melakukan pengumpulan data, melakukan observasi dan mengumpulkan *dataset* proposal skripsi yang akan digunakan sebagai alat pembandingan.

2. Perencanaan

Fase ini dimaksudkan untuk menggambarkan tugas-tugas teknis yang harus dilakukan, risiko-risiko yang mungkin muncul, sumber daya yang dibutuhkan, produk-produk kerja yang harus dihasilkan dan jadwal-jadwal

kerja. Tahap ini akan dilaksanakan dengan cara membuat jadwal pelaksanaan atau *schedule* pembuatan *software* dan melakukan pemilihan terhadap data yang diperlukan.

3. Permodelan

Fase ini dimaksudkan untuk membuat gambaran besar seperti apa yang akan dibuat nantinya baik dengan membuat sketsa-sketsa supaya lebih dapat memahami masalah yang sesungguhnya. Tahap ini direalisasikan, dimana representasi *interface* di desain menggunakan *Balsamiq Mockups 3*, arsitektur *software* dibuat dengan *UML Diagram*.

4. Konstruksi

Fase ini merupakan kegiatan yang menggabungkan kode dan pengujian yang sangat dibutuhkan untuk menemukan kekeliruan-kekeliruan dalam kode program yang dihasilkan sebelumnya. Tahap ini direalisasikan ke dalam bentuk kode program dalam bahasa *PHP* dengan menggunakan bantuan *tools Notepad++* , selanjutnya permodelan yang telah dibuat diuji dengan pengujian *blackbox* dan pengujian *whitebox* terhadap sistem tersebut.

5. Penyerahan sistem/perangkat lunak ke para pelanggan/ pengguna

Fase ini merupakan penyerahan perangkat lunak kepada pelanggan yang kemudian akan mengevaluasi produk yang disajikan dan akan memberikan umpan balik berdasarkan evaluasi tersebut. Tahap ini direalisasikan

Pemilihan model *waterfall* dalam pembuatan aplikasi untuk mengukur tingkat kemiripan dokumen ini didasari oleh berbagai pertimbangan salah satunya alur dalam pengerjaan proyek yang jelas sehingga proses pembangunan aplikasi dapat dilakukan secara disiplin dan dapat menghindari kesalahan estimasi waktu pengembangan aplikasi. Hal ini cukup penting karena waktu pengembangan sistem ini tergolong singkat.

1.6. Sistematika Penulisan

Penulisan laporan dilakukan secara sistematis menggunakan berapa bab. Gambaran umum mengenai isi laporan penelitian secara keseluruhan akan diuraikan secara singkat dengan sistematika sebagai berikut.

BAB I PENDAHULUAN

Bab ini berisikan latar belakang permasalahan yang dibahas dalam penelitian, perumusan masalah, tujuan dan manfaat penelitian, metodologi penelitian dan sistematika penulisan laporan penelitian.

BAB II LANDASAN TEORI

Bab ini berisikan teori – teori dasar yang digunakan sebagai landasan dan referensi dalam penelitian. Teori tersebut mencakup mengenai teori mengenai sistem temu kembali (*information retrieval*), *similarity document*, Algoritma *Rabin Karp*, Algoritma *TF-IDF*, Algoritma *Vector Space Model*, *Unifield Modeling Language (UML)* yang digunakan dalam penelitian.

BAB III ANALISIS DAN PERANCANGAN SISTEM

Bab ini berisikan analisis kebutuhan sistem, permodelan UML, desain arsitektur sistem dan desain antarmuka sistem pada sistem untuk mengukur kemiripan dokumen yang akan dibangun.

BAB IV IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab ini berisikan pembahasan dari proses pembuatan sistem, seperti hasil penerapan desain dalam sistem dan juga pengujian fungsionalitas pada sistem.

BAB V KESIMPULAN DAN SARAN

Bab ini berisikan kesimpulan yang dapat diambil dari penelitian yang telah dilakukan serta saran-saran yang dapat berguna untuk penelitian dan pengembangan sistem selanjutnya.