

Bab I

Pendahuluan

1.1 Latar Belakang

Plagiarisme merupakan penyimpangan yang banyak terjadi salah satunya pada bidang pendidikan di tingkat perguruan tinggi. Mahasiswa yang sedang mengerjakan tugas akhir, proposal penelitian, maupun laporan sementara praktikum berpotensi melakukan plagiarisme secara sengaja maupun tidak sengaja. Plagiarisme menurut Kamus Besar Bahasa Indonesia (KBBI) berarti menggambar atau menulis garis-garis gambaran atau tulisan yang telah tersedia (dengan menempelkan kertas kosong pada gambar atau tulisan yang akan ditiru), mencontoh atau meniru tulisan atau pekerjaan orang lain, mencuri karangan orang lain dan mengakui sebagai karangan sendiri, mengutip karangan orang lain tanpa seizin penulisnya (KBBI, 1997).

Plagiarisme atau plagiat adalah suatu perbuatan menjiplak ide, gagasan atau karya orang lain yang selanjutnya diakui sebagai karya sendiri atau menggunakan karya orang lain tanpa menyebutkan sumbernya sehingga menimbulkan asumsi yang salah atau keliru mengenai asal muasal dari suatu ide, gagasan atau karya. Istilah plagiat berasal dari bahasa Inggris *plagiarism* atau *plagiary* serta dalam bahasa Latin *plagiarius* yang artinya penculik atau penjiplak. Jadi plagiarisme atau plagiat adalah tindakan mencuri (gagasan/karya intelektual) orang lain dan mengklaim atau mengumumkannya sebagai miliknya.

Sedangkan menurut Peraturan Menteri Pendidikan Nasional Republik Indonesia No.17 tahun 2010, plagiat merupakan perbuatan secara sengaja atau tidak sengaja dalam memperoleh atau mencoba memperoleh kredit atau nilai untuk suatu karya ilmiah, dengan mengutip sebagian atau seluruh karya dan/atau karya ilmiah pihak lain yang diakui sebagai karya ilmiahnya, tanpa menyatakan sumber secara tepat dan memadai. Pelaku plagiarisme akan dikenakan hukuman dengan pidana penjara masing-masing paling singkat 1 (satu) bulan dan/atau

denda paling sedikit Rp 1.000.000,00 (satu juta rupiah), atau pidana penjara paling lama 7 (tujuh) tahun dan/atau denda paling banyak Rp 5.000.000.000,00 (lima miliar rupiah). Plagiarisme sangat dilarang oleh undang-undang pemerintah Republik Indonesia karena hal ini sangat merugikan dan mengandung hak cipta seseorang.

Pada Prodi Pendidikan Informatika terdapat matakuliah praktikum disetiap semesternya yang mana matakuliah ini bertujuan untuk menerapkan secara langsung segala materi dan teori yang telah mahasiswa peroleh dari hasil belajar di kelas. Kemudian setelah melakukan praktik, mahasiswa diminta untuk menuliskan hasil yang didapat dalam sebuah laporan sementara. Tentunya laporan setiap mahasiswa tidak akan sama karena berdasarkan hasil praktik yang diperoleh masing masing mahasiswa. Sampai saat ini masih belum diketahui apakah seluruh mahasiswa melakukan penulisan laporan sementara praktikum dengan jujur tanpa adanya plagiarisme. Plagiarisme merupakan tindakan yang tidak benar karena jika mahasiswa melakukan tindakan ini, maka proses penyerapan ilmu yang terjadi pada mahasiswa tidak akan maksimal. Selain itu tindakan plagiasi juga melanggar hak cipta dan norma yang berlaku dimasyarakat. Berdasarkan hasil wawancara terhadap salah satu dosen pengampu matakuliah praktikum, yakni ibu Laili Cahyani S.Kom.,M.Kom, beliau menegaskan bahwa belum terdapat sistem yang mampu mendeteksi plagiarisme terhadap berkas laporan sementara praktikum. Oleh karena itu tentu sangat dibutuhkan suatu sistem yang mampu melakukan pendeteksian plagiarisme pada dokumen laporan sementara praktikum. Hal ini akan sangat membantu dosen maupun asisten praktikum dalam pengoreksian dokumen laporan sementara praktikum dengan cepat dan akurat. Pengecekan plagiarisme tersebut dapat diselesaikan dengan menggunakan algoritma Rabin – Karp yang dikombinasikan dengan *Levensthein Distance*.

Algoritma Rabin Karp adalah algoritma pencarian kata yang mencari sebuah pola berupa substring dalam sebuah teks menggunakan hashing. Algoritma ini sangat efektif untuk pencocokan kata dengan pola banyak. Berdasarkan penelitian dari Pandu Pratama Putra dalam jurnalnya menyatakan bahwa dari beberapa pengujian yang telah dilakukan, algoritma Rabin – Karp dapat

menemukan kutipan terpanjang yang sama di antara dua dokumen teks serta mengukur kemiripan dokumen teks. Dan Muhammad Elfan Efendi dalam penelitian pendeteksi plagiarisme menggunakan metode *Levenshtein Distance* menyatakan bahwa Hasil pengujian dari skenario 1 sampai 4 menunjukkan bahwa jarak rata-rata nilai kemiripan penggunaan stemming dan nonstemming tidak terlalu jauh, yaitu 0,26% dan perbedaan rata-rata jarak waktu sekitar 10 detik. Sedangkan hasil skenario 5 dan 6 dalam penggunaan filtering dan tidak filtering menunjukkan perbedaan hasil sebesar 5%. Sedangkan Noviardi Prima Putra dan Sularno dalam penelitiannya mengkombinasikan algoritma Rabin – Karp dengan metode *Synonym Recognition* sehingga dapat menemukan kata yang benar-benar sama atau kata yang sudah dirubah kebentuk kata yang bermakna sama. Pada penelitian yang dilakukan oleh Andry Hery Purba dan Zakarias Situmorang dengan menggunakan algoritma Rabin – Karp dan *Levenshtein Distance* menegaskan bahwa semakin banyak jumlah karakter atau kalimat teks yang hendak diukur tingkat plagiarismenya, maka akan semakin besar atau semakin lama waktu yang dibutuhkan untuk melakukan proses pendeteksian plagiarisme.

Selain itu menurut Brinardi Leonardo dan Seng Hansun pada penelitiannya mengemukakan bahwa algoritma Rabin – Karp lebih baik dibandingkan algoritma Jaro-Winkler *Distance* dalam melakukan pendeteksian plagiarisme. Hal ini dibuktikan dengan melakukan pengujian tingkat plagiarisme terhadap beberapa dokumen. Pengujian dokumen menggunakan algoritma Rabin-Karp menghasilkan rata-rata persentase *similarity* sebesar 51% sedangkan pengujian dokumen menggunakan algoritma Jaro-Winkler Distance menghasilkan rata-rata persentase *similarity* sebesar 35%. Selain itu pada penelitiannya juga menegaskan bahwa algoritma Rabin-Karp mampu melakukan proses pendeteksian plagiarisme dengan waktu yang lebih singkat dibandingkan dengan algoritma Jaro-Winkler. Dan algoritma Rabin-Karp juga lebih mampu mendeteksi plagiarisme pada dokumen dengan ukuran yang sangat besar.

Dewi Rosmala dan Zulfikar Muhammad Risyad pada penelitiannya menegaskan bahwa penggunaan algoritma *Levenshtein Distance* dalam menemukan kesalahan ejaan kata, mampu memberikan akurasi 100% terhadap kata yang dicari. Selain itu pada aplikasi *drugs e-dictionary* yang dikembangkan oleh

Halimah Tus Sadiyah, Muhamad Saad Nurul Ishlah dan Nisa Najwa Rokhmah dengan menggunakan algoritma *Levenshtein Distance* mampu memberikan presisi pencarian kata sebesar 90% hal ini sangat berguna untuk mencari data atau informasi secara akurat. Dan yang terakhir adalah penelitian yang dilakukan oleh Andysah Putera Utama Siahaan yang mengkombinasikan algoritma Rabin-Karp dan *Levenshtein Distance* dan menemukan fakta bahwa kombinasi kedua algoritma tersebut mampu meningkatkan akurasi pendeteksian plagiarisme terhadap suatu dokumen yang dipengaruhi oleh tiga parameter utama yakni n-gram, basis dan modulo.

Berdasarkan hasil wawancara dan dan observasi serta latar belakang permasalahan diatas, maka pada penelitian ini akan dibuat **Sistem Pendeteksi Plagiarisme Berbasis Web Dengan Algoritma Rabin-Karp dan Levenshtein Distance**. Penelitian ini merupakan pengembangan dari penelitian sebelumnya yang berjudul Sistem Informasi Manajemen Laboratorium Pendidikan Informatika Menggunakan *Framework Laravel* oleh Muhamammad Syaifuddin. Dengan adanya sistem pendeteksi plagiarisme ini diharapkan dapat menekan jumlah tindakan plagiarisme yang terjadi. Selain itu sistem ini cukup efektif dan akurat untuk melakukan pendeteksian pada dokumen laporan sementara praktikum. Tentu hal ini akan sangat membantu para dosen dan asisten praktikum yang memiliki keterbatasan sebagai manusia biasa dalam melakukan pendeteksian.

1.2 Tujuan Penelitian & Pengembangan

Penelitian ini bertujuan untuk :

- a. Mengembangkan sistem pendeteksi plagiarisme pada Laporan sementara praktikum prodi pendidikan informatika berbasis web.
- b. Untuk mengetahui akurasi persentase algoritma Rabin-Karp dan Levenshtein Distance dalam pendeteksian plagiarisme.

1.3 Spesifikasi Produk

1.3.1 Fitur yang terdapat pada sistem pendeteksi plagiarisme

- a. Terdapat tombol *upload file* yang berfungsi untuk *meng-upload* berkas dokumen yang hendak di cek tingkat plagiarismenya

- b. Terdapat tombol Cek untuk memulai perintah pengecekan tingkat plagiarisem pada dokumen tersebut
- c. Terdapat Laporan Hasil yang berisi persentase hasil pengecekan plagiarisme pada dokumen tersebut yang disajikan dalam bentuk tabel.

1.3.2 Spesifikasi dari segi hardware dan software

- a. Hardware
 - Processor minimal core i5
 - RAM minimal 4GB
 - Hardisk minimal 500GB
- b. Software
 - Web browser

1.4 Batasan Penelitian Pengembangan

- a. Sistem pendeteksi plagiarisme ini hanya dapat mendeteksi tindakan plagiarisme pada dokumen-dokumen internal yakni dokumen laporan sementara yang sudah terkumpul di Prodi Pendidikan Informatika.
- b. Sistem pendeteksi plagiarisme ini diujikan pada laporan sementara praktikum struktur data.
- c. Dokumen teks yang akan dideteksi adalah file teks digital yang bersifat plain text, yaitu file yang hanya terdiri dari huruf-huruf dan angka-angka saja, tidak mencakup gambar, tabel, dan sejenisnya.
- d. Perbandingan oleh sistem dilakukan dengan membandingkan dua dokumen teks yang telah dikonversi ke format text document (.txt).

1.5 Pentingnya Penelitian Pengembangan

1.5.1 Bagi Mahasiswa :

- a. Untuk mencegah terjadinya tindakan plagiarisme pada laporan sementara.

1.5.2 Bagi Prodi Pendidikan Informatika:

- a. Membantu dosen dalam mendeteksi plagiarisme laporan sementara praktikum pendidikan informatika.
- b. Mampu meningkatkan keakuratan seleksi maupun pemeriksaan plagiarisme pada laporan sementara di prodi Pendidikan Informatika.

