

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Perkembangan teknologi informasi yang semakin pesat membawa dampak baik dan buruk bagi kehidupan manusia, dampak buruk yang ditimbulkan salah satunya adalah masalah plagiarisme (Nurdin dkk., 2019). Kasus Plagiarisme saat ini merupakan isu yang sangat sering terdengar, terutama di kalangan akademisi di Perguruan Tinggi. Tindakan meniru karya ilmiah penulis lain tanpa mencantumkan sumber rujukan atau dengan kata lain mengakui karya tersebut merupakan buah pemikiran sendiri telah berlangsung lama di berbagai perguruan tinggi baik negeri maupun swasta di Indonesia (Swari dkk., 2021). Dengan adanya kebijakan Dikti, terhitung mulai tanggal kelulusan setelah Agustus 2012, karya ilmiah mahasiswa harus diupload. Sedangkan untuk dosen, efektif diimplementasikan untuk usulan kenaikan pangkat dan jabatan dosen mulai tahun 2012. Dengan diuploadnya karya ilmiah yang dihasilkan dosen atau mahasiswa tersebut, akan terlihat apakah karya tersebut terindikasi plagiat atau tidak (Lestari dkk., 2019).

Sejak diberlakukannya Surat Edaran Dirjen-DIKTI No.2050/E/T/2011 tentang Kebijakan Unggah Karya Ilmiah dan Jurnal; Surat Edaran Dirjen-DIKTI No. 152/E/T/2012 tentang Publikasi Karya Ilmiah S1 di Jurnal Nasional, S2 di Jurnal

Nasional Terakreditasi, dan S3 di Jurnal Internasional. Kewajiban publikasi yang sudah diatur bagi dosen dan peneliti diharapkan dapat mendorong peningkatan jumlah dan mutu publikasi ilmiah pada tingkat nasional dan internasional sehingga Indonesia mampu bersaing dengan bangsa-bangsa lain. Kemudian diikuti oleh Surat Edaran Dirjen-DIKTI No. 212/E/T/2012 tentang Panduan Pengelolaan Jurnal Terbitan Berkala Ilmiah Elektronik; dan Surat Edaran Dirjen Risbang-DIKTI No.93/E/SE/XII/2015 tentang Akreditasi Jurnal Ilmiah Secara Elektronik, pemerintah menetapkan kebijakan bahwa mulai 1 April 2016, akreditasi jurnal ilmiah nasional hanya berlaku untuk jurnal yang telah dikelola secara elektronik dengan *platform Open Journal System* (OJS) agar proses penilaian lebih mudah, cepat, akurat dan transparan. Jurnal yang telah diterbitkan secara elektronik kemudian didaftarkan pada database Akreditasi Jurnal Nasional (ARJUNA). Adanya kebijakan tersebut, maka setelah 1 April 2016, para pengelola/penerbit jurnal di Indonesia harus mematuhi segala persyaratan/ketentuan dalam penerbitan e-journal agar terbitannya dapat terakreditasi secara nasional (Nashihuddin, 2020).

Plagiarisme adalah tindakan melanggar hukum dan kode etik, hal ini tentunya menjadi momok dalam perkembangan ilmu penegetahuan. Pelaku plagiarisme bertindak dengan mengambil dan menuliskan karya penulis lain pada karya yang diakui sebagai miliknya tanpa menuliskan asal dari tulisan yang ditulis tersebut (Swari dkk., 2021). Maka dari itu pemerintah, dosen pembimbing, dan pengelola jurnal selalu cermat dan berhati-hati dalam menerima dan memeriksa suatu karya ilmiah, bahkan pemerintah mengeluarkan kebijakan untuk setiap karya ilmiah yang diunggah ke *repository* kampus wajib dicek plagiasinya (Syaharuddin dkk., 2021).

Plagiat merupakan suatu pencurian karangan milik orang lain (Salim dkk., 2021). Kegiatan menjiplak atau mengambil hasil karangan, pendapat, ide dan karya

orang lain dan menjadikannya seolah hal tersebut merupakan karangan, pendapat, ide dan karya sendiri merupakan suatu bentuk plagiarism atau plagiat. Plagiarism atau plagiat merupakan tindakan pidana karena hal ini merupakan sebuah pencurian, yaitu mencuri hak cipta orang lain. Di dunia Pendidikan, pelaku plagiarism akan mendapatkan hukuman berat seperti dikeluarkan dari sekolah atau universitas. Pelaku plagiat disebut sebagai *plagiator*. Menurut (Syaharuddin dkk., 2021) tindakan plagiat disebabkan oleh beberapa faktor, diantaranya: (1) minimnya jiwa literasi mahasiswa untuk membaca banyak referensi atau artikel-artikel yang sudah terbit secara online, (2) terbatasnya aplikasi untuk mengecek plagiat yang terjangkau, karena harus membayar untuk lisensi, (3) mahasiswa tidak memahami prosedur cek plagiat dan cara memperbaiki plagiasinya, (4) saat ini banyak beredar hasil-hasil penelitian yang belum terunggah ke *repository* ataupun *publish* di jurnal secara *online*. Salah satu upaya pencegahan plagiarisme pada karya ilmiah mahasiswa yaitu dengan menerapkan kebijakan *open access* di institusional *repository*. Institusional *repository* adalah sebuah perpustakaan digital yang mampu menampung koleksi-koleksi digital berupa thesis, disertasi, artikel, jurnal dan lain sebagainya. Sedangkan *open access* adalah sebuah kebijakan yang dibuat oleh para ilmuwan professional untuk menerbitkan jurnal-jurnal atau karya-karya ilmiah lain yang dapat dibaca secara gratis tanpa harus melanggannya terlebih dahulu (Zahroh & Wahyudin, 2020).

Kesamaan suatu dokumen dengan dokumen lainnya atau suatu *text* dengan *text* lainnya dapat diketahui dengan menggunakan pendekatan *string metric* yaitu dengan melakukan perbandingan dua *string*. Kedua *string* tersebut dimasukkan ke dalam fungsi matematis tertentu untuk mengetahui jarak antara keduanya (Rouf dkk., 2020). Teknik kesamaan adalah konsep dasar yang sangat banyak digunakan.

Misalnya, kesamaan antara dua subjek yaitu subjek A dan dan subjek B terkait dengan kesamaan ataupun perbedaan di antara keduanya. Menghitung jumlah kemiripan kata adalah tugas umum, namun memberikan peran yang sangat penting dalam berbagai aplikasi *Natural Language Processing* (NLP), seperti mesin pencari, detektor plagiarisme, sistem penjawab pertanyaan, dan lainnya (Vatolin dkk., 2021). Tentu saja menghitung jumlah kemiripan kata sangat bisa digunakan untuk mendeteksi tingkat plagiat pada artikel ilmiah, mengingat karya ilmiah merupakan hasil pemikiran seseorang yang dituangkan dalam bentuk teks. Algoritma yang dapat digunakan untuk mencari kesamaan string dan kemiripan kata yaitu *Jaro-Winkler*, *Hamming*, *Damerau Levenshtein*, dan *Levenshtein Distance*. Algoritma *Jaro-Winkler* dipilih karena sangat cocok dan sesuai dalam memberikan hasil terbaik pada pencocokan dua *string* pendek. Penggunaan algoritma *Jaro-Winkler* telah dilakukan oleh peneliti sebelumnya yaitu, koreksi otomatis dan saran ejaan, fitur, mendeteksi dokumen plagiarisme, *stemming* kata kompensasi tidak standar bahasa Inggris (Yulianto & Nurhasanah, 2021).

Algoritma *Jaro-Winkler* juga bisa diterapkan untuk mendeteksi kesalahan ejaan. Salah mengeja merupakan fenomena umum, terutama saat berhubungan dengan *internet*. Ejaan yang buruk membuat seseorang tampak kurang cerdas dan kurang kredibel daripada yang sebenarnya. Kesalahan ejaan tidak hanya merusak reputasi profesional seseorang, tetapi juga menghabiskan banyak uang dalam penjualan dan bisnis. Kesalahan ejaan dalam kemasan medis bisa berakibat fatal dan bisa berakhir dengan merenggut nyawa seseorang. Jadi memeriksa kesalahan ejaan dan memperbaikinya adalah layanan yang sangat dibutuhkan di setiap Bahasa (Ahamed dkk., 2021).

Algoritma *Jaro-Winkler* menghitung tingkat kesamaan antara dua kata. Semakin tinggi nilai *Jaro-Winkler* maka semakin tinggi tingkat kemiripannya. Algoritma ini telah digunakan oleh Nur dalam penelitiannya yang menyimpulkan bahwa analisis sentimen dengan perbaikan kata menggunakan algoritma *Jaro-Winkler Distance* lebih baik dari pada tanpa perbaikan kata yang ditunjukkan dari nilai *precision* dan *accuracy* (Nur, 2021). Algoritma *Jaro-Winkler Distance* digunakan untuk mendeteksi plagiarisme dokumen skripsi dengan cara melakukan perbandingan antara dokumen asli dan dokumen uji sehingga dapat mendeteksi *similarity* pada dokumen baik (Aksari dkk., 2020).

Menurut (Ahamed dkk., 2021) di dunia *online*, terutama di *platform* media sosial kebanyakan dari kita menulis tanpa banyak memperhatikan ejaan dan tata bahasa yang benar. Kesalahan ejaan jauh lebih besar dalam proporsi ketika datang ke bahasa Bangla. Dalam penelitiannya menyajikan metode untuk deteksi kesalahan dan koreksi dalam ejaan kata-kata Bangla. Sistem dapat mendeteksi kata Bangla yang salah eja dan menyediakan dua layanan yang menyarankan ejaan yang benar untuk kata tersebut. Serta mengoreksi kata tersebut dengan menggunakan algoritma *Norvig* tetapi menggunakan probabilitas kata-kata yang disiapkan untuk saran dan koreksi dengan jarak menggunakan *Jaro-Winkler*. Pekerjaan sebelumnya yang dilakukan di bidang ini untuk bahasa Bangla sangat lambat atau kurang akurat. Dengan adanya sistem ini hasil mencapai akurasi 97% ketika dievaluasi dengan 1000 kata Bangla.

Menurut (Finansyah dkk., 2022) semakin banyaknya data yang disimpan dalam bentuk teks membuat pengolahan data teks lebih sulit. Hal ini juga menyebabkan masalah pada algoritma *text preprocessing*, salah satunya adalah

ketika dua teks identik, tetapi dianggap berbeda oleh algoritma. Sehingga perlu dilakukan normalisasi teks untuk mendapatkan bentuk standar dari kata-kata dalam bahasa tertentu. Dengan data tekstual yang semakin meningkat pada *platform* media sosial mengukur kesamaan semantik teks pendek menjadi lebih penting, aplikasi *Natural Language Processing* mendeteksi *parafrase* yang mengukur kesetaraan semantik dari dua bagian teks (Shahmohammadi dkk., 2021).

Meninjau apa yang disebut jurnal predator mungkin membuang-buang waktu dan tenaga. Sebagai salah satu dari banyak kekurangan, jurnal predator tidak menjamin pengarsipan dan akses jangka panjang ke isinya. Publikasi mereka menghasilkan pembaca terbatas, jarang dikutip dan akhirnya mungkin hilang. Meskipun memberikan umpan balik pada manuskrip yang dikirimkan ke jurnal predator mungkin membantu penulis meningkatkan pekerjaan mereka, meninjau manuskrip yang hampir tidak memiliki dampak ilmiah tidak mungkin memajukan sains. Juga, ada laporan anekdot bahwa *peer review* di jurnal predator tidak memiliki integritas ilmiah. Untuk menyoroti beberapa masalah jurnal predator, seorang peneliti mengirimkan manuskrip yang cacat secara sengaja dan pengulas melihat permasalahan pada karya ilmiah tersebut, sementara editor menerima karya ilmiah itu. Sama seperti penerbitan penelitian di jurnal predator cenderung membuang-buang sumber daya, meninjau untuk outlet ini harus dilihat dalam cahaya kritis yang sama, dan upaya diperlukan untuk menghentikan praktik ini (Severin dkk., 2021).

Menurut pandangan pendapat tokoh ahli di atas, bisa ditarik kesimpulan adalah hasil sebuah artikel ilmiah yang berbentuk sebuah pemaparan tulisan hasil dari sebuah penelitian yaitu membahas suatu permasalahan dalam suatu bidang tertentu dengan menggunakan kaidah-kaidah yang berlaku. Akan tetapi seiring

dengan berjalannya waktu perkembangan informasi serta teknologi, tak hanya memberikan efek positif tetapi memberikan efek negatif. Hal ini diakibatkan banyaknya bertebaran sebuah informasi yang sering dimanfaatkan sebagai rujukan di dalam pembuatan artikel ilmiah. Sehingga tak jarang ditemukan kecurangan seperti meniru ide orang lain maupun menjiplak artikel ilmiah yang juga diistilahkan dengan *plagiarisme*.

Berdasarkan latar belakang di atas maka dilakukan penelitian dengan judul Akurasi dalam Mengidentifikasi Tingkat *Similarity* pada Artikel Ilmiah Menggunakan *Algorithma Jaro Winkler*.

1.2 Perumusan Masalah

Berdasarkan uraian latar belakang di atas, dapat dirumuskan permasalahan penelitian adalah:

1. Bagaimana menerapkan *Algorithma Jaro Winkler* untuk mengidentifikasi tingkat *similarity* pada artikel ilmiah?
2. Bagaimana pengujian aplikasi deteksi *plagiarism* yang diterapkan di STKIP Rokania?

1.3 Batasan Masalah

Pembatasan permasalahan dalam membuat perangkat lunak adalah sebagai berikut:

1. Penelitian ini hanya menguji teks judul dan abstrak artikel ilmiah berbahasa Indonesia dan bukan berbentuk berkas.
2. *Sample* data perbandingan berupa judul dan abstrak yang digunakan sifatnya *real* yaitu dari E-Jurnal STKIP Rokania.
3. Hasil kemiripan data uji dengan data pembanding disimpan dalam *database* hanya dalam bentuk presentase kemiripan.

1.4 Tujuan Penelitian

Adapun tujuan penelitian yang dilakukan adalah sebagai berikut:

1. Membangun sistem pendeteksi tingkat similarity pada artikel ilmiah berbasis web terintegrasi dengan E-Jurnal STKIP Rokania.
2. Menerapkan sistem pendeteksi tingkat similarity pada artikel ilmiah di STKIP Rokania.
3. Mengetahui tingkat akurasi kegiatan plagiarisme pada dokumen artikel ilmiah.

1.5 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian yang dilakukan yaitu:

1. Sebagai sebuah media alat bantu peneliti dalam mengoreksi artikel ilmiahnya.
2. Untuk mengetahui tingkat kemiripan artikel ilmiah milik sendiri dengan artikel ilmiah milik orang lain.
3. Hasil dari penelitian ini diharapkan dapat digunakan sebagai tambahan pengetahuan bagi pembaca atau bagi peneliti lain, serta dapat menjadi referensi untuk penelitian sejenis.
4. Hasil akhir dari penelitian ini berupa *website online* yang bisa diakses dan digunakan oleh masyarakat umum.

1.6 Sistematika Penulisan

Penulisan dan penyusunan penelitian ini dilakukan secara sistematis agar mempermudah pembaca dalam memahami isi penelitian ini, adapun sistematika penulisannya adalah sebagai berikut:

BAB I : PENDAHULUAN

Dalam bab ini berisi tentang latar belakang dari masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

BAB II : LANDASAN TEORI

Dalam bab ini dijelaskan mengenai landasan teori yang menjadi landasan untuk penelitian, serta membahas pengertian dari metode yang diangkat dan sistem yang akan digunakan untuk pemodelan perancangan sistem.

BAB III: METODOLOGI PENELITIAN

BAB ini berisi tentang analisa dan perancangan yang akan membahas mengenai sumber data yang digunakan, metode pengumpulan data dan alat bantu penelitian.

BAB IV : ANALISA DAN PERANCANGAN

Bab ini berisi tentang pengujian data ke dalam sistem yang dirancang untuk menyelesaikan permasalahan yang diangkat, kebutuhan sistem, implementasi sistem, kelemahan dan kelebihan pada sistem yang dibangun.

BAB V : IMPLEMENTASI DAN HASIL

Dalam bab ini berisi implementasi sistem yang telah diterapkan dan hasil pengujian terhadap sistem yang baru beserta saran-saran untuk kemajuan ataupun pengembangan dimasa yang akan datang.

BAB VI: KESIMPULAN DAN SARAN

Bab ini membahas hasil pengujian yang dirangkum dalam bentuk kesimpulan yang meliputi kelebihan dan kekurangan dari penelitian ini yang menjadi saran dan masukan yang membangun.