

1. Pendahuluan

Kemajuan teknologi saat ini dapat dikatakan sebagai salah satu penyebab dari semakin berkembangnya cara dari seseorang untuk menyelesaikan masalahnya. Tidak hanya hal itu, seseorang pada dasarnya menginginkan sesuatu hal berjalan dengan cepat atau instan sehingga tidak jarang mereka menghalalkan segala cara untuk mencapai keinginannya. Hal ini juga terjadi dalam pembuatan sesuatu benda atau karya cipta. Berdasarkan tindak kecurangan tersebut dikenal istilah plagiat atau plagiarisme untuk mendapatkan hasil kerja yang lebih cepat tanpa perlu mengeluarkan banyak usaha dan tenaga.

Plagiarisme adalah suatu kegiatan yang mengambil ide dari orang lain dan mengakuinya sebagai ide miliknya [1]. Saat ini tindakan seperti ini sering dilakukan apalagi dengan dukungan teknologi seperti internet dimana kita dapat mencari berbagai macam sumber bacaan. Oleh karena itu diperlukan suatu sistem yang dapat membandingkan dokumen asli dengan dokumen yang merupakan hasil plagiarisme. Pada penelitian ini akan menggunakan pendekatan *Text Alignment* untuk mendeteksi plagiarisme dengan bantuan metode *Source Retrieval*. *Text Alignment* merupakan suatu pendekatan yang digunakan untuk membantu mengidentifikasi semua bagian text yang telah digunakan dan digunakan kembali pada dua buah dokumen. *Source retrieval* sendiri merupakan inti dari kegiatan pendeteksi plagiat misalkan saat diberi sebuah dokumen *suspicious* atau mencurigakan dan sebuah *web search engine* dan tugas yang akan dijalankan adalah untuk mengambil dokumen sumber yang mana teks nya telah digunakan kembali. Sistem ini dimulai dengan proses *source retrieval* untuk memproses dokumen *suspicious* yang nantinya akan menghasilkan potongan atau fragment dari dokumen sumber yang akan diproses menggunakan pendekatan *text alignment* untuk dihitung score plagiarismenya.

1.1. Latar Belakang

Plagiarisme atau menjiplak dapat diartikan menggunakan kembali karya orang lain seolah – olah karya tersebut merupakan karya miliknya sendiri. Plagiarisme teks mungkin adalah salah satu bentuk plagiarisme tertua yang pernah ada sampai saat ini dan masih sulit untuk diidentifikasi dalam praktiknya saat ini. Oleh karena itu banyak penelitian yang sudah dilakukan untuk mendeteksi plagiarisme secara otomatis [7]. Dalam hal ini PAN @CLEF rutin mengadakan kompetisi dalam hal mendeteksi plagiat sejak tahun 2009. Dan semenjak kompetisi ini diadakan, telah banyak kerangka kerja yang diajukan untuk mengatasi kasus plagiarisme. Salah satunya adalah menggunakan pendekatan *Text Alignment* serta metode *Source Retrieval*. Pada penelitian ini akan menggunakan pendekatan *Text Alignment* untuk mendeteksi plagiarisme dengan bantuan metode *Source Retrieval* yang akan mengumpulkan fragmen dari dokumen sumber yang nantinya akan dicocokkan dengan *suspicious document* oleh *Text Alignment* dan dihitung tingkat kecocokannya. Metode ini dipilih karena menurut jurnal yang telah diterbitkan oleh pelaksana PAN @CLEF kedua metode tersebut dapat digabungkan yaitu dengan menggunakan *source retrieval* kita dapat mencari dokumen yang memiliki kesamaan dengan dokumen yang dicurigai plagiat dengan berbagai dokumen yang terdapat di internet dengan bantuan Chatnoir API. Selain itu kita juga bisa mendapat fragmen dari dokumen sumber yang diduga mirip dengan dokumen yang dicurigai untuk nanti di bandingkan dengan pendekatan *text alignment* untuk mendapatkan hasil atau score tingkat kemiripan dokumen tersebut.

1.2. Topik dan Batasannya

Untuk mendeteksi tindak plagiarisme diperlukan suatu sistem yang mampu mendeteksi kesamaan pada dokumen yang dianggap plagiat (*suspicious document*). Dalam penelitian ini akan dilakukan pengujian sistem pendeteksi plagiat menggunakan metode *source retrieval* dan pendekatan *text alignment*. Adapun rumusan masalah yang diangkat penulis berdasarkan topik diatas adalah bagaimana mengimplementasikan metode *source retrieval* dan *text alignment* dalam sistem deteksi plagiat serta bagaimana performansi yang dihasilkan oleh sistem tersebut nantinya.

Dalam penelitian ini hanya akan dilakukan pengujian untuk menemukan source dokumen dari dokumen yang diduga plagiat serta mengumpulkan fragment dokumennya untuk diproses dan dihitung kesamaannya menggunakan *text alignment*. Output yang dihasilkan dalam penelitian ini hanya berupa score tingkat plagiarisme antara dokumen mencurigakan dan potongan dokumen dari hasil pencarian *source retrieval*.

1.3. Tujuan

Berdasarkan latar belakang serta rumusan masalah diatas maka tujuan dari penelitian ini adalah untuk mengimplementasikan metode *source retrieval* dan *text alignment* dalam sistem deteksi plagiat dan menganalisis hasil dari sistem yang telah dibuat. Selain itu, penelitian ini juga dimaksudkan untuk mengetahui performansi dari sistem yang telah dibuat.