

## 1. PENDAHULUAN

### 1.1 Latar Belakang

Universitas Kristen Petra memiliki berbagai macam jurusan dimana salah satunya adalah Teknik Informatika. Setiap tahunnya teknik informatika pastinya akan menghasilkan sejumlah mahasiswa yang akan maju ke tahap penyusunan skripsi. Seiring dengan peningkatan jumlah mahasiswa yang akan mengajukan skripsi, penting bagi teknik informatika untuk memastikan bahwa topik yang diajukan oleh mahasiswa tidak memiliki kesamaan dengan skripsi yang telah dibuat oleh mahasiswa sebelumnya. Hal ini penting karena perbedaan topik skripsi merupakan kriteria penting bagi Universitas Kristen Petra.

Hingga saat ini pengecekan topik skripsi oleh mahasiswa masih dilakukan dengan menggunakan *website* pra skripsi yang hanya mencocokkan judul skripsi saja, dimana metode yang digunakan yaitu *string matching*. Oleh karena penggunaan kriteria hanya merupakan judul skripsi maka fitur tersebut kurang dapat memberikan jawaban secara akurat. Hal ini dikarenakan terdapat judul skripsi yang diajukan memang berbeda dengan yang sudah dibuat sebelumnya namun isi pembahasan memiliki kemungkinan untuk sama, dan begitu juga sebaliknya. Dimana kriteria yang dibutuhkan untuk pengecekan topik skripsi merupakan judul skripsi, kata kunci, abstrak, latar belakang dan kesimpulan.

Dengan kurangnya keakuratan *website* pra skripsi yang dibuat oleh alumni untuk memberikan jawaban pengecekan topik maka mahasiswa harus bertanya kepada koordinator skripsi apakah topik yang diajukan dapat digunakan. Oleh karena mahasiswa harus bertanya kepada koordinator skripsi, pertanyaan yang diterima koordinator skripsi menjadi banyak. Apabila mahasiswa memiliki topik yang mirip dosen juga harus melakukan analisa kembali apakah tujuan atau metode yang diajukan berbeda. Hal ini dapat menyebabkan proses dosen untuk memberikan jawaban kepada mahasiswa menjadi tidak efisien.

Dengan permasalahan yang ada maka dibutuhkan sebuah sistem pengecekan yang dapat memenuhi kebutuhan dari teknik informatika secara praktis dan efisien. Sistem pengecekan yang akan dibuat akan menggunakan metode TF-IDF (*Term Frequency - Inverse Document Frequency*) dan akan dibantu menggunakan teknik *Phrase Recognition* yang dapat meningkatkan keakuratan kesamaan topik skripsi. Penelitian telah dilakukan sebelumnya oleh (Lindang et al, 2022) dimana mereka menggunakan metode TF-IDF dan *Cosine Similarity* untuk penentuan kemiripan antar skripsi pada perpustakaan. Hasil penelitian tersebut menunjukkan

bahwa metode yang digunakan memiliki tingkat akurasi pengujian rata rata sebesar 97%. Metode TF-IDF dapat digunakan untuk membantu sistem pengecekan kesamaan karena TF-IDF dapat memberikan hasil seberapa relevan masing masing kata dalam suatu dokumen. Pengecekan kesamaan topik skripsi tidak bergantung pada urutan kata, melainkan seberapa relevan sebuah kata dalam kriteria seperti tujuan, metode, serta obyek skripsi. Oleh karena penelitian yang telah dilakukan sebelumnya, peneliti ingin menggabungkan metode TF-IDF dengan teknik *Phrase Recognition* untuk meningkatkan hasil yang akan didapat dikarenakan *Phrase Recognition* dapat membantu dalam mendapatkan *multi* kata. Sehingga dengan adanya sistem pengecekan ini diharapkan dapat membantu mahasiswa serta dosen teknik informatika untuk menjalankan kegiatan skripsi dengan baik dan efisien.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dibuat, maka dapat di rumuskan permasalahan sebagai berikut :

1. Apakah aplikasi sistem pengecekan kemiripan dapat membantu Teknik Informatika dalam mengatasi kemiripan topik skripsi yang diajukan mahasiswa?
2. Apakah implementasi TF-IDF dan *Phrase Recognition* dapat memberikan jawaban yang *valid* terhadap pengecekan kesamaan topik skripsi?

## 1.3 Tujuan

Tujuan dari skripsi ini adalah untuk merancang dan membuat sistem pengecekan kesamaan topik skripsi mahasiswa pada *website* pra skripsi untuk membantu mahasiswa serta dosen pada saat mahasiswa ingin mengajukan topik skripsi, sehingga pengecekan kesamaan topik skripsi dapat dilakukan secara cepat, otomatis dan *real time* tanpa harus menunggu jawaban dari dosen.

## 1.4 Ruang Lingkup

Ruang lingkup dibatasi pada:

1. Aplikasi dibuat pada *website*.
2. *Code editor* menggunakan Visual Studio Code.
3. Pembuatan desain *prototype* menggunakan Figma.
4. Bahasa pemrograman menggunakan Javascript dengan *framework* Vue.js.

5. Sistem API dibuat menggunakan Node.js.
6. Pengujian sistem API menggunakan aplikasi Postman.
7. *Database* yang digunakan adalah PostgreSQL.
8. Pengujian pada aplikasi yang dibuat akan menggunakan *survey* kepada koordinator skripsi serta mahasiswa teknik informatika.
9. Sumber data skripsi mahasiswa yaitu dari :
  - a. Dewey *Online Catalog* Universitas Kristen Petra
  - b. Jurnal *Infra* Universitas Kristen Petra
  - c. Skripsi yang sedang berjalan
10. Data yang akan diambil dari sumber data skripsi yaitu :
  - a. Nama mahasiswa
  - b. *File* PDF skripsi
11. Data *file* PDF skripsi akan di proses sebagai berikut :
  - a. PDF yang sudah diambil akan dilakukan proses pembacaan
  - b. Proses pembacaan PDF akan dibantu menggunakan *library* PDF.js
  - c. *Text* yang didapatkan dari proses pembacaan PDF akan dilanjutkan ke proses *preprocessing data* sebelum masuk ke tahap selanjutnya
12. Data skripsi yang sedang berjalan terdiri dari skripsi yang telah lulus sidang proposal skripsi.
13. Kriteria pengecekan kesamaan topik skripsi mahasiswa :
  - a. Judul skripsi
  - b. Kata kunci
  - c. Abstrak
  - d. Latar Belakang
  - e. Kesimpulan

14. Abstrak dan latar belakang digunakan maksimal 2 paragraf inti sebagai pengecekan.
15. Paragraf inti abstrak dan latar belakang akan dilakukan pengamatan pada 5 skripsi teknik informatika yang telah dipublikasikan dengan topik yang identik.
16. Fitur tambahan merupakan penjawab otomatis *Frequently Asked Questions*.
17. Pertanyaan dan jawaban dari *Frequently Asked Questions* sudah ditetapkan oleh dosen koordinator skripsi.
18. Fitur *website* dosen :
  - a. Melihat daftar judul yang di *submit* oleh mahasiswa.
  - b. Melihat urutan kemiripan judul proposal dengan skripsi yang sudah dibuat.
  - c. Mengatur daftar pertanyaan dan jawaban untuk *Frequently Asked Questions*.
19. Fitur *website* mahasiswa :
  - a. *Input* judul skripsi
  - b. *Input* kata kunci
  - c. *Input* deskripsi singkat mengenai topik dengan maksimal 200 kata
  - d. Melihat daftar skripsi yang memiliki kemiripan dengan *ranking*.
  - e. Melakukan *submit* judul skripsi.
20. Hasil dari aplikasi yang dibuat adalah *website* yang dapat menunjukkan tingkat kemiripan topik skripsi dengan urutan teratas yang paling mirip diikuti dengan persentase kemiripan.

## 1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah menghasilkan suatu aplikasi berbasis *web* yang dapat membantu mahasiswa teknik informatika dalam melakukan pengecekan kesamaan topik skripsi sehingga dapat mendapatkan informasi secara efisien dan mengantisipasi terjadinya duplikasi skripsi. Penelitian ini juga dapat membantu dosen teknik informatika agar tidak perlu lagi mengecek kesamaan topik skripsi mahasiswa dengan cara konvensional yang sangat memakan waktu.

## 1.6 Metodologi Penelitian

Langkah - langkah dalam pengerjaan skripsi :

1. Studi Literatur
  - 1.1. Metode TF-IDF (*Term Frequency - Inverse Document Frequency*)
  - 1.2. Teknik *Phrase Recognition*

- 1.3. *Cosine Similarity*
  - 1.4. *Framework Vue.js*
  - 1.5. *Node.js*
  - 1.6. *Database PostgreSQL*
2. Perencanaan dan Pembuatan Aplikasi
    - 2.1. Melakukan analisa pada abstrak dan latar belakang pada skripsi yang telah dibuat.
    - 2.2. Pembuatan alur kerja sistem
    - 2.3. Pembuatan desain *Entity Relationship Diagram*
    - 2.4. Pembuatan *database*
    - 2.5. Pembuatan servis sistem API dan aplikasi
  3. Pengujian Aplikasi
    - 3.1. Melakukan pengujian terhadap sistem aplikasi apakah dapat berjalan dengan baik dan efisien.
    - 3.2. Melakukan perbaikan apabila terdapat kesalahan atau penambahan fitur.
  4. *Deployment* pada *server* Universitas Kristen Petra
  5. Pengambilan Kesimpulan
    - 5.1. Pembuatan hasil analisa dari pengujian sistem.
    - 5.2. Pembuatan kesimpulan mengenai hasil penelitian yang sudah dilakukan.
  6. Pembuatan Laporan
    - 6.1. Pembuatan laporan dari hasil yang telah diperoleh.

## **1.7 Sistematika Penulisan**

Penulisan laporan skripsi ini dibagi menjadi beberapa bab, yaitu:

**BAB I : PENDAHULUAN**

Bab ini berisikan judul, latar belakang, perumusan masalah, ruang lingkup, tujuan skripsi, dan metodologi penelitian yang akan digunakan dalam skripsi ini.

**BAB II : LANDASAN TEORI**

Bab ini berisikan teori - teori yang akan digunakan dan diterapkan dalam skripsi ini.

**BAB III : ANALISA DAN DESAIN SISTEM**

Bab ini berisikan analisa dan desain sistem yang akan dibuat.

**BAB IV : IMPLEMENTASI SISTEM**

Bab ini berisikan tentang implementasi sistem berdasarkan desain sistem.

**BAB V : PENGUJIAN SISTEM**

Bab ini berisikan pengujian sistem dan hasil pengujiannya dari implementasi yang telah dilakukan pada Bab IV.

**BAB VI : KESIMPULAN DAN SARAN**

Bab ini berisikan kesimpulan yang dapat diambil dari hasil yang dicapai dan saran untuk penelitian selanjutnya.