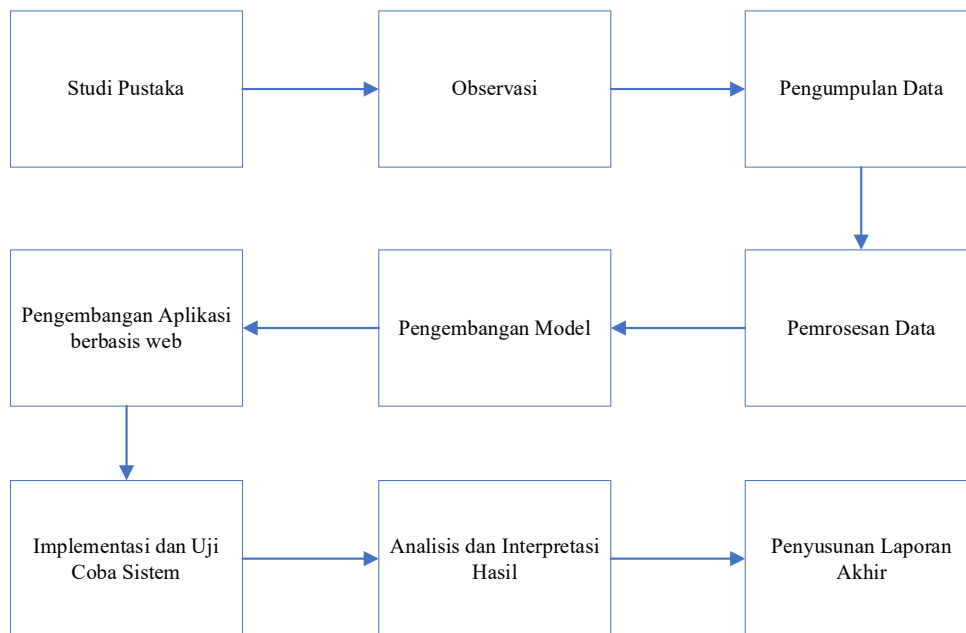


## BAB III METODE PENELITIAN

### 3.1 Alur Penelitian



*Gambar 3. 1 Alur Penelitian*

Penjelasan dari alur penelitian sebagai berikut:

#### 1. Studi Pustaka

Proses studi pustaka dilakukan untuk mencari referensi terkait penelitian yang pernah dilakukan sebelumnya, termasuk jurnal, artikel ilmiah, buku, dan sumber lainnya yang relevan. Tahap ini bertujuan untuk memahami permasalahan, metodologi, dan hasil penelitian yang dapat menjadi dasar untuk penelitian yang akan dilakukan.

#### 2. Observasi

Observasi dilakukan untuk mengidentifikasi kebutuhan dan masalah nyata yang ada pada sistem atau kasus tertentu. Dalam penelitian ini, observasi membantu memahami pola publikasi dan kebutuhan akan prediksi publikasi penelitian.

### 3. Pengumpulan Data

Tahap ini mencakup pengumpulan dataset dari artikel penelitian, termasuk atribut seperti judul, abstrak, penulis, tahun, dan lainnya. Data ini diambil dari Goodwood Publishing yang telah terbit dan terindeks di google scholar.

### 4. Pemrosesan Data

Data yang terkumpul kemudian diproses untuk memastikan kualitasnya. Proses ini mencakup pembersihan data, penghapusan data duplikat, pengisian data yang hilang, dan transformasi data teks menjadi format terstruktur yang siap digunakan oleh algoritma.

### 5. Pengembangan Model

Algoritma Random Forest diterapkan pada data yang telah diproses untuk membangun model prediksi. Proses ini melibatkan pembagian dataset menjadi data latih dan data uji, pelatihan model, serta evaluasi performa model menggunakan metrik yang sesuai.

### 6. Pengembangan Aplikasi Berbasis Web

Pada tahap ini, aplikasi berbasis web dikembangkan untuk memfasilitasi pengguna dalam menggunakan model prediksi. Aplikasi ini didesain untuk menerima input seperti judul, abstrak dan kata kunci dan memberikan hasil prediksi secara real-time.

### 7. Implementasi dan Uji Coba Sistem

Aplikasi yang telah dikembangkan diuji untuk memastikan fungsionalitas dan keakuratan prediksi. Proses ini juga mencakup pengujian performa sistem serta pengumpulan umpan balik dari pengguna untuk meningkatkan aplikasi.

#### 8. Analisis dan Interpretasi Hasil

Hasil prediksi dianalisis untuk melihat pola atau tren yang dapat digunakan sebagai bahan pengambilan keputusan. Tahap ini juga mencakup evaluasi dampak hasil penelitian terhadap pengembangan ilmu pengetahuan dan kebutuhan pengguna.

#### 9. Penyusunan Laporan Akhir

Penelitian diakhiri dengan penyusunan laporan akhir yang mencakup seluruh proses, hasil, dan rekomendasi penelitian. Laporan ini disusun untuk memenuhi persyaratan akademik dan disampaikan kepada pihak-pihak terkait.

### 3.2 Jenis Penelitian

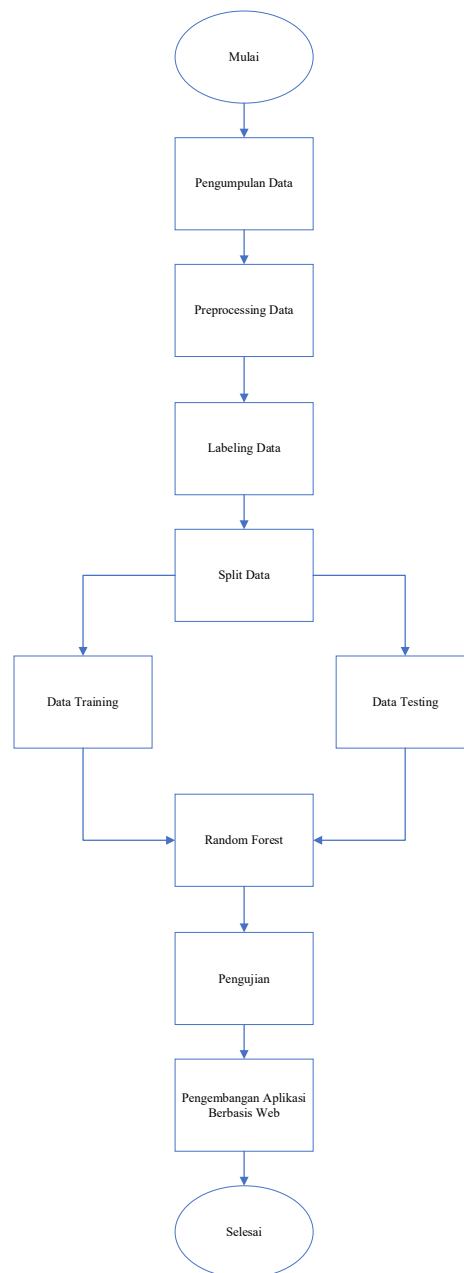
Penelitian kuantitatif eksperimen adalah metode penelitian yang bertujuan untuk menguji hubungan sebab-akibat antara variabel dengan melakukan intervensi atau perlakuan tertentu terhadap objek penelitian (Santoso & Madiistriyatno, 2021). Dalam penelitian ini, data dikumpulkan, dianalisis secara numerik, dan diuji menggunakan metode statistik untuk mengukur dampak dari suatu perlakuan. Pendekatan ini sering digunakan dalam bidang ilmu komputer, sains, dan teknik, termasuk dalam pengujian model machine learning seperti Random Forest untuk mengevaluasi performanya dalam melakukan prediksi berdasarkan dataset yang tersedia.

### 3.3 Sumber Data

Data yang digunakan dalam penelitian ini diperoleh dari Goodwood Publishing, yang merupakan penerbit jurnal ilmiah yang menerbitkan berbagai artikel penelitian dalam berbagai disiplin ilmu. Untuk mengumpulkan data artikel yang relevan, penelitian ini menggunakan aplikasi *Publish or Perish*, sebuah perangkat lunak yang memungkinkan pengambilan data bibliografi dari berbagai sumber akademik.

Aplikasi *Publish or Perish* digunakan untuk mengekstrak metadata artikel dari Goodwood Publishing, termasuk judul, kata kunci, abstrak, jumlah kata, tanggal publikasi, jumlah sitasi, dan bidang studi. Dalam penelitian ini data yang digunakan meliputi judul dan abstrak. ini kemudian diolah untuk dianalisis menggunakan algoritma Random Forest, dengan tujuan mengklasifikasikan potensi publikasi artikel berdasarkan pola yang telah teridentifikasi dalam dataset.

### 3.4 Diagram Alur Penelitian



*Gambar 3. 2 Diagram Alur Penelitian*

a. Pengumpulan Data

Data dikumpulkan dari Goodwood Publishing, mencakup judul artikel, kata kunci, abstrak, tahun publikasi, serta atribut lainnya. Data yang terkumpul akan digunakan sebagai dataset utama penelitian ini.

b. Preprocessing Data

Tahapan preprocessing dilakukan untuk membersihkan dan mempersiapkan data agar siap diproses:

- 1) Cleaning Data: Membersihkan data dari duplikasi atau informasi yang tidak relevan.
- 2) Tokenisasi: Memecah data teks menjadi unit kata.
- 3) Stopword Removal: Menghapus kata-kata yang tidak signifikan (contoh: "yang", "dan").
- 4) Stemming: Mengubah kata ke bentuk dasar (contoh: “pengembangan” → “kembang”).
- 5) TF-IDF: Memberikan bobot pada kata-kata dalam data teks menggunakan teknik TF-IDF untuk mengetahui pentingnya kata pada tiap dokumen.

c. Labeling Data

- 1) Data diberi label berdasarkan kategori tertentu, seperti topik penelitian, jumlah sitasi, atau kemungkinan menarik minat pembaca.
- 2) Proses ini memastikan bahwa data siap digunakan sebagai input dalam model klasifikasi.

d. Split Data

- 1) Dataset dibagi menjadi data training dan data testing dengan perbandingan, misalnya 80:20.
- 2) Data training digunakan untuk melatih model Random Forest, sedangkan data testing digunakan untuk menguji akurasi model.

e. Implementasi Random Forest

Tahapan penerapan algoritma Random Forest mencakup:

- 1) Bootstrap Sampling: Membuat sampel acak dari dataset.
- 2) Random Feature Selection: Memilih atribut secara acak untuk membentuk pohon keputusan.

- 3) Model Training: Menumbuhkan banyak pohon keputusan tanpa pruning.
- 4) Majority Vote: Hasil klasifikasi ditentukan berdasarkan suara mayoritas dari seluruh pohon yang terbentuk.

f. Pengujian Model

Model diuji menggunakan Confusion Matrix untuk mendapatkan hasil evaluasi. Metrik yang digunakan antara lain:

- a) Akurasi: Tingkat ketepatan model.
- b) Precision: Ketepatan dalam mengklasifikasikan kelas tertentu.
- c) Recall: Kemampuan model menemukan semua data yang relevan.
- d) F1-Score: Harmoni antara Precision dan Recall.

g. Pengembangan Aplikasi Berbasis Web

- 1) Model Random Forest diintegrasikan ke dalam aplikasi berbasis web.
- 2) Aplikasi ini memiliki fungsi utama untuk mengklasifikasikan topik relevan dan perkiraan publikasi penelitian yang berpotensi menarik minat pembaca.

### 3.5 Metode Analisis Data

1. Langkah-Langkah Pemrosesan Data

- a. Preprocessing Data: Melakukan pembersihan data (data cleaning) dan ekstraksi fitur, seperti transformasi teks menggunakan TF-IDF untuk mengubah data teks menjadi representasi numerik (Rabbani, Safitri, Rahmadhani, & Anam, 2023).
- b. Pembagian Data: Membagi dataset menjadi data pelatihan dan data uji (80% untuk pelatihan dan 20% untuk pengujian).

2. Pelatihan dan Pengujian Model

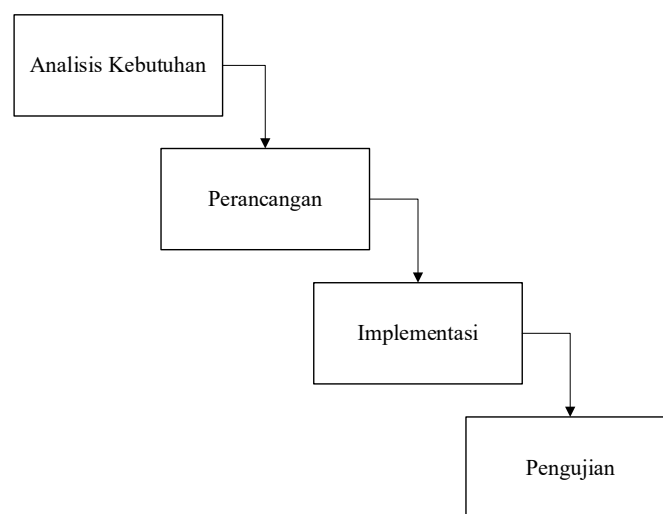
- a. Melatih algoritma Random Forest menggunakan data pelatihan.
- b. Menguji model yang telah dilatih pada data uji untuk mengukur kinerja klasifikasi.

### 3. Evaluasi Kinerja Model

Metrik evaluasi yang digunakan meliputi akurasi, precision, recall, F1-score, waktu pemrosesan, serta penggunaan sumber daya.

### 3.6 Metode Pengembangan Aplikasi

*System Development Life Cycle (SDLC)* dengan model *Waterfall* menyediakan kerangka kerja yang sistematis untuk mengembangkan aplikasi berbasis web:



Gambar 3. 3 Alur perancangan SDLC Waterfall

#### 1. Analisis Kebutuhan

Tahap ini dilakukan untuk mengidentifikasi kebutuhan fungsional dan non-fungsional sistem berdasarkan tujuan penelitian. Kebutuhan utama yang harus dipenuhi oleh sistem meliputi:

- Sistem dapat memproses input berupa judul dan abstrak artikel.
- Sistem dapat melakukan klasifikasi kategori artikel secara otomatis.
- Sistem dapat memberikan rekomendasi jurnal sesuai hasil klasifikasi.
- Sistem menyediakan laporan hasil klasifikasi yang dapat diunduh.
- Sistem memiliki autentikasi pengguna untuk admin dan user.

Table 3. 1 Fucntional Requirement

| Kode | Nama Fitur                | Deskripsi                                                                  |
|------|---------------------------|----------------------------------------------------------------------------|
| FR1  | Sistem Klasifikasi        | Memproses input judul dan abstrak untuk menentukan kategori artikel.       |
| FR2  | Laporan Hasil Klasifikasi | Menyediakan laporan hasil klasifikasi yang dapat diunduh dalam format PDF. |
| FR3  | User Authentication       | Fitur login untuk membedakan akses pengguna dan admin.                     |
| FR4  | Rekomendasi Jurnal        | Memberikan rekomendasi jurnal sesuai hasil klasifikasi.                    |
| FR5  | Manajemen Data Admin      | Admin dapat melihat, menghapus, dan mengelola data hasil klasifikasi.      |

## 2. Perancangan

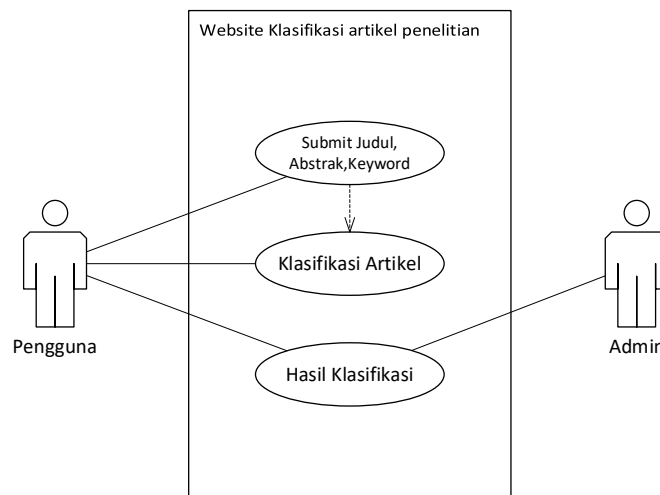
### a. Use Case

Use Case Diagram menggambarkan interaksi antara sistem dan aktor yang terlibat. Dalam sistem ini terdapat dua aktor utama, yaitu:

- User: Pengguna umum yang dapat melakukan input judul dan abstrak artikel untuk diklasifikasikan.
- Admin: Pengelola sistem yang memiliki akses untuk melihat, mengelola hasil klasifikasi, dan mengelola data pengguna.

Gambar 3.4 merupakan *Use Case Diagram* yang menggambarkan interaksi antara aktor dan sistem dalam *Website Klasifikasi Artikel Penelitian*. Dalam sistem ini, terdapat dua aktor utama, yaitu Pengguna dan Admin. Pengguna memiliki peran untuk memasukkan judul, abstrak, dan keyword ke dalam sistem melalui fitur Submit Judul, Abstrak, Keyword. Setelah data dikirimkan, sistem akan memproses Klasifikasi Artikel, yang bertujuan untuk menentukan kategori atau klasifikasi dari artikel yang dimasukkan. Hasil dari proses ini dapat diakses melalui fitur Hasil Klasifikasi, yang dapat dilihat oleh baik pengguna maupun admin.

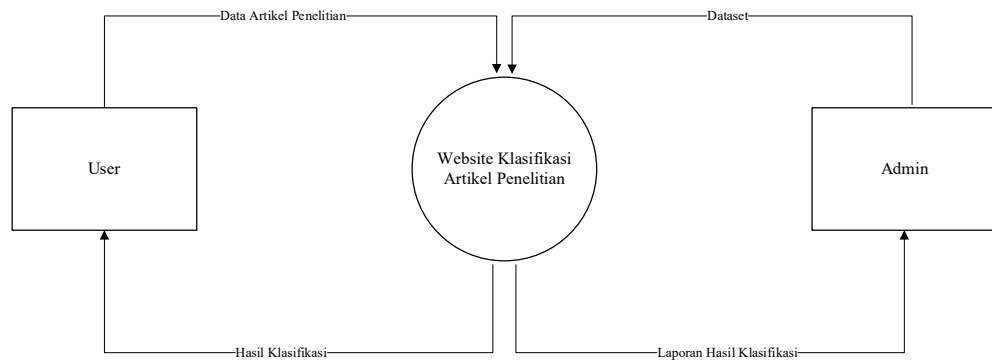
Admin juga memiliki akses terhadap hasil klasifikasi untuk keperluan analisis dan pengelolaan data. Diagram ini menunjukkan bagaimana sistem bekerja dalam membantu pengguna mengklasifikasikan artikel penelitian berdasarkan data yang diberikan.



Gambar 3. 4 Use case

#### b. Context Diagram

Context Diagram menunjukkan hubungan aliran data antara sistem dan aktor eksternal. Gambar 3.5 merupakan *Context Diagram* yang menggambarkan hubungan antara sistem *Website Klasifikasi Artikel Penelitian* dengan aktor-aktor yang berinteraksi dengannya. Dalam diagram ini, terdapat dua aktor utama, yaitu User dan Admin. User berperan dalam menginput Data Artikel Penelitian ke dalam sistem untuk diproses. Setelah melalui proses klasifikasi, sistem akan menghasilkan Hasil Klasifikasi, yang kemudian dikembalikan kepada user. Sementara itu, Admin bertanggung jawab dalam mengelola sistem dengan menyediakan Dataset sebagai referensi untuk klasifikasi serta menerima Laporan Hasil Klasifikasi dari sistem. Diagram ini secara keseluruhan menggambarkan bagaimana data mengalir antara user, admin, dan sistem dalam proses klasifikasi artikel penelitian.



Gambar 3. 5 Context Diagram

### c. Rancangan Antarmuka Aplikasi

Antarmuka aplikasi didesain sederhana dan user-friendly, mencakup:

#### — Halaman Login dan Registrasi

**Prediction Journal**
Login
Register

**Login**

Username

Password

Don't have an account? Register [here](#)

Gambar 3. 6 Rancangan Antarmuka Halaman Login

**Prediction Journal**
Login
Register

**Register**

Username

Password

Already have an account? Login [here](#)

Gambar 3. 7 Rancangan Antarmuka Halaman Register

— Halaman Input Judul dan Abstrak

The screenshot shows a web application interface for 'Prediction Journal'. At the top, there is a header with the site name and 'Login' and 'Register' links. The main content area is titled 'Classification' and contains three input fields: 'Article Title', 'Abstract', and a 'Predict' button.

Gambar 3. 8 Rancangan Antarmuka Halaman Classification

— Halaman Hasil Klasifikasi dan Rekomendasi

The screenshot shows a web application interface for 'Prediction Journal'. At the top, there is a header with the site name and 'Login' and 'Register' links. The main content area is titled 'Classification Results' and contains a table with columns: Title, Abstract, Category, and Recommended Journals. There are also 'Download PDF' and 'Delete' buttons.

Gambar 3. 9 Rancangan Antarmuka Halaman Classification Result

### 3. Implementasi

Tahap ini adalah proses pengembangan perangkat lunak berdasarkan desain yang telah dibuat, termasuk penulisan kode dan pembuatan database.

### 4. Pengujian

Mengimplementasikan model klasifikasi ke dalam aplikasi berbasis web.