

## BAB 2

### Landasan Teori

#### 2.1 Tinjauan terhadap penelitian terkait

Dibawah ini merupakan hasil penelitian terkait yang menjadi referensi dalam pengembangan penelitian untuk keperluan analisis.

1. Sumarni Adi dalam penelitiannya menjelaskan bahwa setiap mahasiswa akhir harus mengusulkan judul skripsi ke bagian admisi program studinya. Salah satu syarat judul skripsi dikatakan layak adalah judul tersebut belum pernah digunakan oleh mahasiswa lainnya. Namun setiap semester begitu banyak mahasiswa yang mengajukan judul skripsi, sehingga pihak program studi kewalahan jika harus mencocokkan judul skripsi yang masuk dengan judul skripsi sebelumnya karena membutuhkan energi dan waktu yg cukup banyak. Untuk itu penelitian ini mengusulkan algoritma Rabin Karp untuk mendeteksi kemiripan judul skripsi dengan memberikan nilai kemiripan pada judul tersebut. (adi, 2018)
2. Sugiono, Herwin , Hamdani, dan Erlin dalam penelitiannya menjelaskan bahwa tindakan copy paste dokumen teks sering terjadi dalam penulisan karya ilmiah tanpa memberikan kredit kepada yang mempunyai document teks tersebut. Tujuan dari penelitian ini adalah untuk membangun sebuah aplikasi yang mampu mendeteksi tingkat kesamaan dokumen teks. (sugiono, herwin, hamdani, & erlin, 2018)
3. Ranti Eka Putri dan Andysah Putera Utama Siahaan dalam penelitiannya menjelaskan bahwa dokument tidak selalu memiliki

*content* yang sama. Namun kesamaan dokument sering serjadi di dunia dalam penulisan karya ilmiah, beberapa kesamaan terjadi karena kebetulan, dan beberapa terjadi karena unsur kesengajaan, pada dokumen yang memiliki sedikit content, ini bisa di periksa oleh mata. Namun pada dokumen yang memiliki ratusan baris dan halaman tidak mungkin. Untuk mengantisipasi ini dibutuhkan suatu cara yang dapat menganalisis plagiarisme. Penelitian ini mengusulkan algoritma rabin karb yang merupakan salah satu dari teknik teks processing. (Putri & Siahaan, 2017)

**Table 2.1** Penelitian Terdahulu

No	Judul	penulis	isi	algoritma
1	penerapan algoritma rabin karp untuk mendeteksi kemiripan judul skripsi	adi, sumarni	Implementasi algoritma pada tingkat kesamaan pada judul skripsi	Rabin Karb
2	Aplikasi Pendeteksi Tingkat Kesamaan Dokumen Teks:Algoritma	( sugiono, herwin, hamdani, & erlin, 2018 )	Membandingkan kedua algoritma pada document teks untuk 1 pebanding untuk mencari yang lebih unggul,	Rabin Karb dan Winnowing

	Rabin Karp Vs Winnowing		winnowing lebih akurat	
3	Examination of Document Similarity Using Rabin-Karp Algorithm	(Putri & Siahaan, 2017)	Implementasi algoritma pada tingkat kesamaan pada document teks	Rabin Karb

Pada penelitian ini yang dilakukan adalah implementasi pada tiga hal kemiripan judul, kemiripan text, serta kemiripan document text dan membandingkan kedua algoritma winnowing dan rabin karb pada tingkat kesamaan dan waktu yang dibutuhkan untuk melakukan scanning.

## 2.2 Similarity Index

*Similarity index* (index kemiripan) merupakan estimasi seberapa banyak pharasa dari kalimat yang di indetifikasi dari document yang dipublish sebelumnya. (chung, lee, lee, park, & kim, 2017)

## 2.3 Plagiarisme

Plagiarisme adalah sebuah tindakan mengambil seluruh atau sebagian tanpa melakukan *parafphrase* dan mengakuinya sebagai karangan (pendapatnya) tanpa memberikan kutipan dari tulisan aslinya yang disalahgunakan. Plagiarisme dapat terjadi secara disengaja maupun tidak disengaja. kedua alasan ini tetap dianggap sebagai plagiarisme jika pada karya ilmiah tersebut terdapat kesamaan tanpa melakukan sitasi dan perubahan teks asli menggunakan kata kata sendiri.

Paraphrase merupakan pengungkapan kembali dengan cara lain namun dengan Bahasa yang sama.

(Muhammad Abdan Shadiq, 2019).

#### **2.4 Jenis Plagiarisme**

Plagiarisme dibagi menjadi 3 tipe yaitu berdasarkan motivasi melakukannya, ditinjau dari cara melakukannya dan self-plagiarisme.

##### **Tipe plagiarisme pertama motivasi melakukan**

Tiga jenis plagiarisme dari kategori ini adalah intentional, unintentional, dan inadvertent. Ketiga jenis ini memiliki unsur kesengajaan.

##### **Tipe plagiarisme kedua ditinjau dari cara melakukan**

Tiga jenis plagiarisme dari kategori ini adalah pathwriting, inappropriate praraphasing, dan summaries. Ketiga jenis ini tidak menggunakan kaidah pengutipan dan paraphrase secara baik.

##### **Tipe Plagiarisme ketiga Self-Plagiarisme**

Tiga jenis plagiarisme dari kategori ini adalah *text recycling*, *redundant and duplicate publication* dan *data fragmentation*. Bentuk plagiarisme yang dilakukan terhadap karya sendiri yang sudah dibuat sebelumnya.

(Shadiqi, 2019)

## 2.5 Algoritma

Algoritma dalam computer digunakan untuk menyelesaikan permasalahan yang ditulis secara runtutan atau terurut dalam set intruksi yang di translate ke dalam Bahasa program pada computer.

### 2.5.1 Rabin Karb

Rabin karb adalah salah satu algoritma pencarian string yang dikembangkan oleh Michael O. Rabin dan Richard M. Karb pada tahun 1987 yang menggunakan teknik hasing untuk menemukan pattern di dalam string teks. Karakteristik menggunakan N-gram dan hasing, penerapannya dilakukan setelah tahapan preprocessing.

#### **Tahapan pada algoritma rabin karb :**

1. N-gram : pembentukan rangkaian Ngram dengan nilai k
2. Rolling hash : pembentukan nilai hashing dari setiap teks document yang dipecah
3. Matching : pencocokan string

(Filcha & Hayaty, 2019)

Pada penelitian ini algoritma rabin karb akan digunakan pada text scanner dan title scanner untuk dibandingkan tingkat akurasi , waktu excecution yang dibutuhkan untuk melakukan score similarity dengan algoritma rabin karb.

### 2.5.2 Winnowing

Winnowing merupakan algoritma yang digunakan untuk melakukan proses pengecekan kesamaan kata (document fingerprinting) untuk mengidentifikasi plagiarisme (penjiplakan). Secara keseluruhan.

## Tahapan pada algorithm winnowing

1. Ngram : Pembentukan rangkaian Ngram
2. Rolling hash : Membentuk nilai hashing dari setiap teks document yang dipecah
3. Window : pembagian karakter dalam window
4. Fingerprint : mencari nilai hashing terkecil setiap window

(Chandra, M, & S.Naga, 2020)

1. input berupa string berupa “suka main basket”.
2. *process case folding* berubah menjadi “sukamainbasket”.
3. *process* pemberian ngram berdasarkan k sebesar 7 (teks akan dipecah 7 huruf) “[sukamai, ukamain, kamainb, amainba, mainbas, ainbask, inbaske, nbasket]”.
4. Process rolling hash pemberian nilai hash dari hasil n gram “[2244815, 2239810, 2058523, 1933337, 2088675, 1927857, 2061261, 2103296 ]”.

(sugiono, herwin, hamdani, & erlin, 2018)

Pada penelitian ini algoritma winnowing akan digunakan pada text scanner , document scanner dan title scanner untuk dibandingkan tingkat akurasi , waktu excecution yang dibutuhkan untuk melakukan score similarity dengan algoritma rabin karb

## 2.6 Web Scrapping Data

Web scrapping adalah teknik untuk mendapatkan informasi dari website secara otomatis tanpa harus menyalinnya secara manual. Tujuan dari web scrapper adalah untuk mencari informasi tertentu dan kemudian mengumpulkannya dalam web yang baru. Web scraping berfokus dalam mendapatkan data dengan cara pengambilan dan ekstraksi.

(ayani, Pratiwi, & Muhardi, 2019)

Pada penelitian ini teknik web scraping digunakan untuk mengumpulkan informasi dari web open jurnal untuk mengumpulkan judul – judul penelitian dan dapat digunakan pada fitur aplikasi scan judul skripsi secara temporar.

## 2.7 Arsitektur Program

Berikut merupakan landasan teori aritektur program yang akan dibangun.

### 2.7.1 PHP

PHP (*PHP: Hypertext Preprocessong*) merupakan salah satu Bahasa pemrograman yang berjalan di sisi server atau server side yang bersifat open source, PHP dibuat oleh Danish canidian. PHP memiliki beberapa paradigma pemrograman yaitu: multi-paradigm, functional, procedural, dan object oriented .

(anggraini, 2021, p. 22)

Pada penelitian ini penulis menggunakan Bahasa pemrograman php dianggap bahwa penulis lebih familiar Bahasa PHP dibanding bahasa pemrograman lainnya dan PHP masih cocok untuk membangun aplikasi berbasis web serta memiliki dukungan pembuatan library terpisah menggunakan composer. Bahasa Php digunakan pada sisi backend apps

## 2.7.2 Javascript

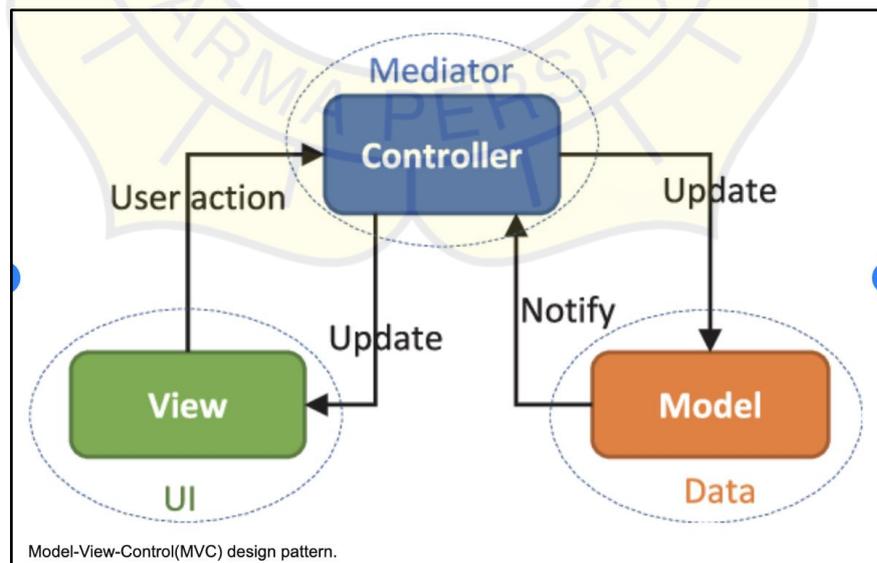
Javascript (js) termasuk Bahasa pemrograman interpreted, javascript merupakan Bahasa paling populer dalam web programming, javascript bisa berjalan di sisi client (browser) dan dapat berjalan di sisi server menggunakan Nodejs. Standar untuk javascript adalah *ecmascript language specification*.

(anggraini, 2021, p. 20)

penelitian ini javascript digunakan untuk pendukung pada frontend apps, bukan murni javascript tapi menggunakan library jquery dan metode ajax untuk digunakan pada text scanner dan template admin

## 2.7.3 MVC

Model view controller (MVC) merupakan pola arsitektur yang umumnya digunakan pada web base aplikasi. MVC mempunyai tiga lapisan utama yaitu, model, view, dan controller. Banyak developer menggunakan MVC sebagai *standar design pattern*.



**Gambar 2.1.** Model View Controller Design Pattern

### **1. Model**

Model class digunakan untuk implementasi logic pada data, di dalam class ini digunakan untuk membuat method insert atau update data ke dalam database

### **2. View**

View digunakan untuk tampilan dari aplikasi, tampilan digunakan untuk interaksi dengan aplikasi

### **3. Controller**

Controller class digunakan untuk menerima user request dan meresponnya melalui controller class dan method yang dipilih user dan mengirimnya kepada pengguna.

Pola MVC menyediakan fasilitas, pengembangan parallel. Ini berarti bahwa setiap lapisan aplikasi independen satu sama lain, tiga pengembang dapat bekerja pada aplikasi tunggal secara bersamaan, satu pengembang akan bekerja pada logika input pengguna, satu pengembang lainnya mengerjakan logika antarmuka pengguna, dan pengembang lainnya akan bekerja pada logika bisnis (model) pada saat yang sama.

(Majeed & Rauf, 2018)

Pada penelitian ini pattern MVC diterapkan pada system aplikasi yang dikembangkan.

#### 2.7.4 ENV

*Environment variable* adalah *dynamic name value* digunakan untuk menjalankan process computer yang memiliki File ber-ekstensi .env memungkinkan untuk di modifikasi. untuk keperluan development.

(Corporation, IBM International Business Machine, 2020)

Salah satu Contoh penggunaan file ber-ekstensi .env adalah ada pada kerangka kerja php yaitu laravel dan codeigniter 4 untuk membuat konfigurasi menggunakan file .env . pada penelitian ini penulis juga menggunakan file ber-ekstensi .env untuk meletakkan variable konfigurasi kedalamnya.

#### **Kode Program 2.1** contoh penggunaan .env pada konfigurasi program computer

```
1. # konfigurasi Path here
2. APP_NAME=orion-scanner
3. APP_HOST=http://localhost/
4.
5. # konfigurasi Database here
6. DB_TYPE=mysql
7. DB_HOST=localhost
8. DB_PORT=3306
9. DB_NAME=orion-scanner
10. DB_USERNAME=root
11. DB_PASSWORD=
```

Dengan meletakkan konfigurasi tersebut maka bisa dipanggil pada php dengan variable global `$_ENV`.

#### **Kode Program 2.2** contoh pemanggilan file .env pada php

```

1. # Database
2. $system["DB_TYPE"] = $_ENV["DB_TYPE"] ? $_ENV["DB_TYPE"] :
   'mysql';
3. $system["DB_HOST"] = $_ENV["DB_HOST"] ? $_ENV["DB_HOST"] :
   'http://localhost/';
4. $system["DB_PORT"] = $_ENV["DB_PORT"] ? $_ENV["DB_PORT"] :
   '3306';
5. $system["DB_NAME"] = $_ENV["DB_NAME"] ? $_ENV["DB_NAME"] : '';
6. $system["DB_USERNAME"] = $_ENV["DB_USERNAME"] ?
   $_ENV["DB_USERNAME"] : 'root';
7. $system["DB_PASSWORD"] = $_ENV["DB_PASSWORD"] ?
   $_ENV["DB_PASSWORD"] : '';

```

### 2.7.5 Composer



**Gambar 2.2.** Composer Dependency manager untuk PHP

Composer merupakan aplikasi package manager untuk Bahasa pemrograman PHP yang format *standar dependency* untuk mengelola *library library* yang diperlukan. Composer dikembangkan oleh Nils Aderman dan Jordi

Boggiano dan release pertama pada tahun 1 Maret 2012. Repository utama composer bernama packagist.

(composer publisher)

Pada penelitian ini composer digunakan untuk pemanggilan library yang dibuat secara terpisah dengan program aplikasi. Dengan tujuan membuat rapih dalam management program. Cara menggunakannya adalah melalui perintah command prompt seperti contoh kode dibawah ini.

### **Kode Program 2.3** contoh penggunaan composer

```
1. PS D:\4.WebDevelopment\2021\orion-scanner> composer require
nagara/metode-skripshit
```

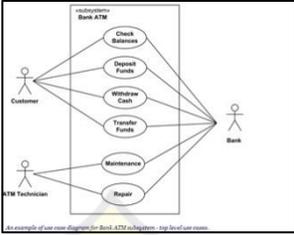
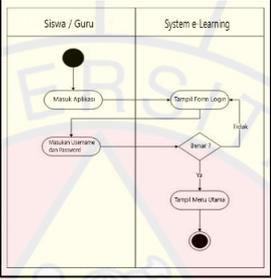
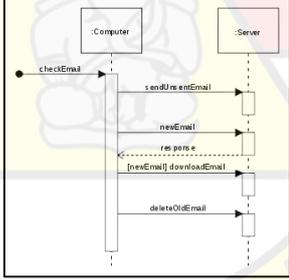
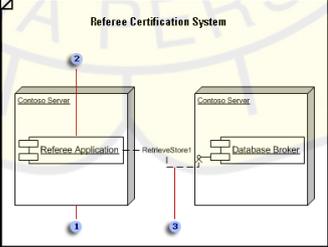
Kode berikut akan mendownload library yang dibuat terpisah pada aplikasi ke dalam aplikasi yang sedang di kembangkan ke dalam directory vendor/nagara

### **2.8 UML Diagram**

UML (Unified Modeling language) merupakan salah satu tools model untuk merancang pengembangan software berbasis object oriented. UML memiliki beragam diagram, beberapa diantaranya yaitu sebagai berikut, use case diagram, activity diagram , sequence diagram, class diagram , component diagram, dan object diagram, namun pada penulisan ini tidak semua diagram digunakan, yang digunakan yakni use case diagram , sequence diagram dan activity diagram.

Sapitri (2021)

**Table 2.2** tabel uml diagram

Nama	Gambar	fungsi
Use case diagram	 <p>The diagram shows a system boundary labeled 'Bank ATM' with several use cases: 'Check Balance', 'Deposit Funds', 'Withdraw Cash', 'Transfer Funds', 'Maintenance', and 'Repair'. Two actors, 'Customer' and 'ATM Technician', are connected to these use cases. 'Customer' is connected to 'Check Balance', 'Deposit Funds', 'Withdraw Cash', and 'Transfer Funds'. 'ATM Technician' is connected to 'Maintenance' and 'Repair'.</p>	<p>Untuk mendeskripsikan sebuah interaksi satu atau lebih actor dengan sistem informasi yang buat.</p>
Activity Diagram	 <p>The diagram is divided into two swimlanes: 'Siswa / Guru' and 'System e-Learning'. In the 'Siswa / Guru' swimlane, the flow starts at a start node, goes to 'Masuk aplikasi', then 'Masuk laman dashboard', and finally 'Masuk laman dashboard'. In the 'System e-Learning' swimlane, the flow starts at a start node, goes to 'Sampul Form Login', then a decision diamond 'Salah?'. If 'Ya' (Yes), it goes to 'Masuk laman dashboard'. If 'Tidak' (No), it goes to 'Sampul Form Login' again. The flow ends at an end node.</p>	<p>Digunakan untuk menggambarkan workflow aktivitas dari sebuah sistem</p>
Sequence Diagram	 <p>The diagram shows two lifelines: ':Computer' and ':Server'. The sequence of messages is: 'checkEmail' (self-call on Computer), 'sendNewEmail' (Computer to Server), 'newEmail' (Server to Computer), 'parse [newEmail] downloadEmail' (Server to Computer), and 'deleteOldEmail' (Computer to Server).</p>	<p>Digunakan untuk menggambarkan interaksi antar objek.</p>
Deployment diagram	 <p>The diagram shows a 'Referee Certification System' node connected to two 'Contoso Server' nodes. The left 'Contoso Server' contains a 'Referee Application' component. The right 'Contoso Server' contains a 'Database Broker' component. A dependency arrow labeled 'RetrieveStores!' points from the 'Referee Application' to the 'Database Broker'.</p>	<p>Digunakan untuk menggambarkan arsitektur sistem.</p>

## 2.9 Agile

Agile merupakan metode pemodelan perangkat lunak yang mengutamakan fleksibilitas terhadap perubahan-perubahan yang terjadi selama proses

pembangunan perangkat lunak. Bahkan perubahan dapat dilakukan pada saat fase terakhir pembangunan perangkat lunak. Salah satu Agile methods adalah Extreme Programming atau disebut juga dengan XP. XP digunakan untuk mengatasi masalah requirements yang tidak jelas dan sering diubah-ubah (vague and volatile requirements).

XP adalah metode pengembangan perangkat lunak yang memberikan kesempatan kepada klien untuk menambahkan atau merubah proses bisnis aplikasi selama pembangunan aplikasi berjalan. Dokumentasi system hanya dilakukan pada tahap eksplorasi dan tahap perencanaan. Kemudian apabila terjadi perubahan pada saat pembuatan aplikasi, perancangan system tidak akan diubah, hanya system yang sedang dibangun saja yang berubah. XP tidak memiliki dokumentasi formal yang dapat digunakan sebagai alat ukur bahwa aplikasi sudah selesai dibangun, namun aplikasi dikatakan selesai apabila klien tidak membutuhkan penambahan requirement pada aplikasi.

Adapun tahapan pembangunan aplikasi menggunakan Xp adalah sebagai berikut.

1. Tahap Eksplorasi

Tahapan eksplorasi adalah setiap kebutuhan yang dituliskan oleh klien akan dibuat dalam bentuk modul yang sederhana atau disebut juga dengan user stories.

2. Tahap Perencanaan

Tahapan perencanaan berorientasi kepada tahapan eksplorasi. Tahapan ini akan memperkirakan kebutuhan bisnis, kebutuhan user, dan kebutuhan system.

### 3. Iterasi Pengembangan Sistem

Pada tahapan ini, akan terjadi beberapa kali iterasi, setiap iterasi terdiri dari 3 tahapan yaitu analisis system, desain system, dan pembuatan dan pengujian system.

### 4. Tahap Produksi Akhir

Tahapan ini dimana system sudah siap untuk di release tahapan ini akan melakukan testing terhadap keseluruhan system yang telah dibuat kepada klien

(Rahmi, Sari, & Subatman, 2016)

Pada penelitian ini XP dipilih karena flexible terhadap perubahan yang diminta oleh klien, dalam hal ini klien adalah penulis sendiri.

**Table 2.3** contoh table user stories pada XP

User	Kebutuhan system
Programmer	Programmer adalah orang yang full akses maintenance aplikasi
Admin	Admin adalah orang yang dapat mengelola resource pada aplikasi meliputi menambahkan user baru, menambah file
Pengguna	Pengguna adalah mahasiswa pegawai dan user yang terdaftar sebagai pengguna untuk menggunakan fitur aplikasi



**TEKNOLOGI INFORMASI**

**UNIVERSITAS DARMA PERSADA**