

BAB III

METODOLOGI PENELITIAN

3.1. Analisis Sistem

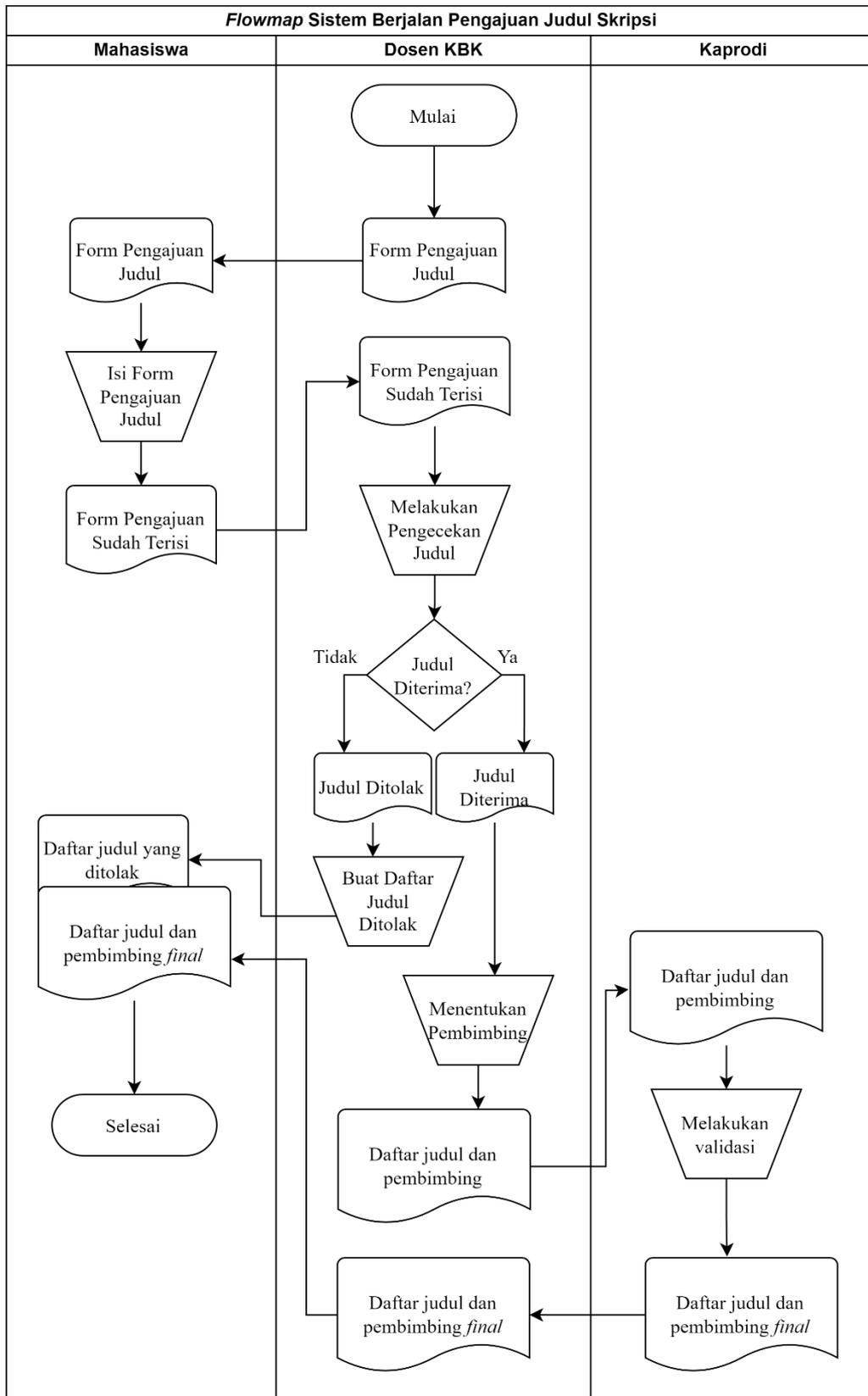
Analisis sistem merupakan tahapan yang pertama kali dilakukan sebelum membangun sistem. Tujuan utama dari melakukan analisis sistem adalah untuk melakukan penelitian atas sistem yang telah ada dan selanjutnya akan merancang sebuah sistem yang baru atau memperbaiki sistem yang telah ada. Analisis sistem berjalan dan sistem yang diajukan adalah sebagai berikut.

3.1.1. Gambaran Umum Sistem yang Berjalan

Pengajuan judul skripsi merupakan tahapan pertama bagi mahasiswa tingkat akhir untuk dapat melakukan penelitian. Salah satu syarat agar judul yang diajukan dapat diterima adalah judul yang akan diajukan belum pernah ada sebelumnya atau tidak sama dengan judul-judul yang telah ada sebelumnya. Pada Program Studi Teknik Informatika Institut Informatika dan Bisnis Darmajaya dalam hal pengajuan judul skripsi masih bersifat manual, yaitu dengan cara mahasiswa datang menemui dosen KBK untuk mengajukan judul skripsi yang akan dikerjakan. Setelah itu dosen KBK akan mencatat judul skripsi yang akan diajukan lalu mengecek serta mengingat apakah judul yang akan diajukan tersebut sudah pernah ada atau belum pernah ada ataupun hanya terdapat sedikit kemiripan dengan judul-judul yang sudah ada sebelumnya (dilakukan secara manual).

3.1.1.1. Bagan Alir (*Flowmap*) Sistem Berjalan

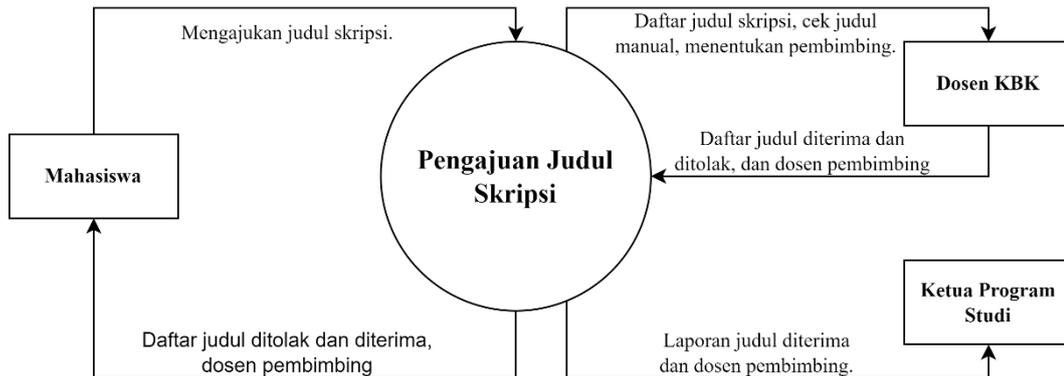
Bagan alir atau *flowmap* merupakan hubungan antara entitas yang terlibat yang menunjukkan arus dari dokumen serta formulir-formulir yang termasuk juga tembusan-tembusannya. *Flowmap* pengajuan judul skripsi pada Program Studi Teknik Informatika Institut Informatika dan Bisnis Darmajaya dapat dilihat pada gambar 3.1.



Gambar 3. 1 *Flowmap* Sistem Berjalan

3.1.1.2. Diagram Konteks Sistem Berjalan

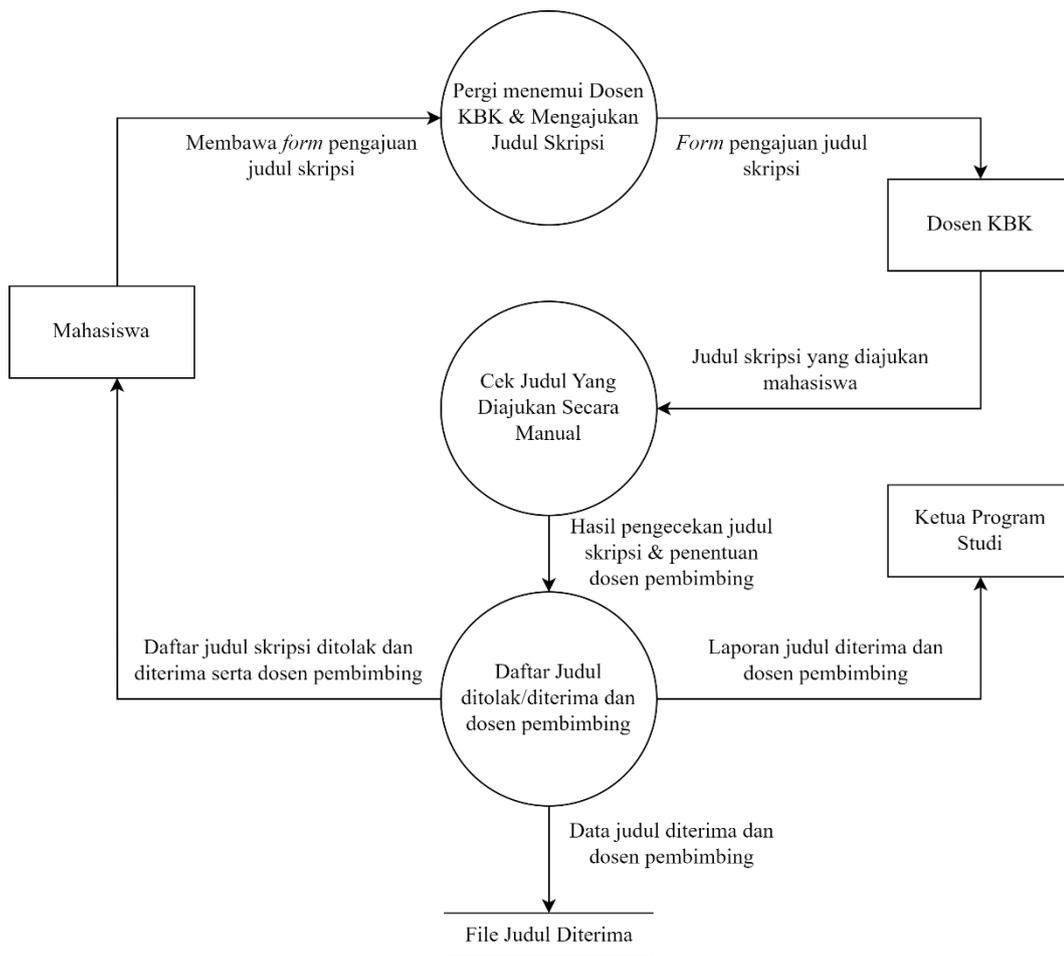
Diagram konteks merupakan diagram tingkatan paling atas yang menjelaskan hubungan sistem secara umum. Diagram konteks dari sistem pengajuan judul skripsi yang sedang berjalan dapat dilihat pada gambar 3.2.



Gambar 3. 2 Diagram Konteks Sistem Berjalan

3.1.1.3. Data Flow Diagram Sistem Berjalan

Data flow diagram merupakan diagram yang menjelaskan lebih mendetail daripada diagram konteks. *Data Flow Diagram* level 1 dari sistem berjalan pengajuan judul skripsi dapat dilihat pada gambar 3.3.



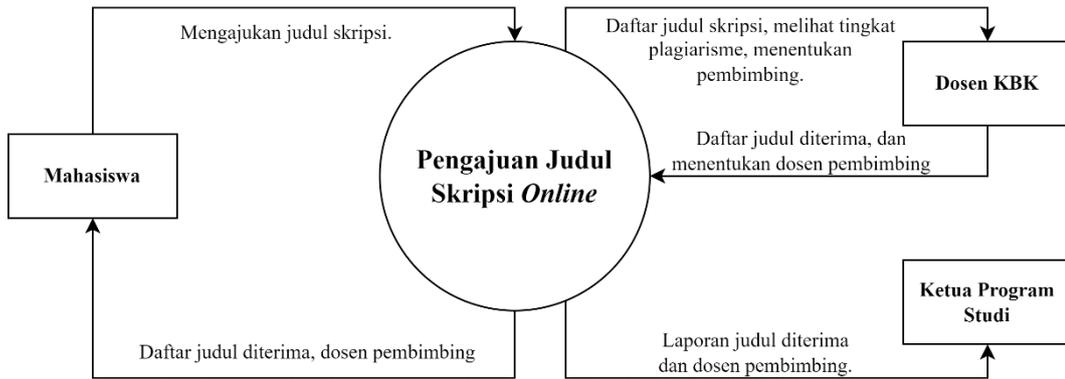
Gambar 3. 3 DFD Level 1 Sistem Berjalan

3.1.2. Gambaran Umum Sistem yang Diajukan

Mendeteksi kemiripan judul skripsi dapat dilakukan dengan menggunakan berbagai metode untuk dapat menghasilkan persentase kemiripan dari judul yang diajukan dengan judul yang sudah ada sebelumnya. Adapun tujuan utama dari mendeteksi kemiripan judul tersebut adalah untuk dapat menentukan persentase kemiripan judul satu dengan judul yang lainnya.

3.1.2.1. Diagram Konteks Sistem Yang Diajukan

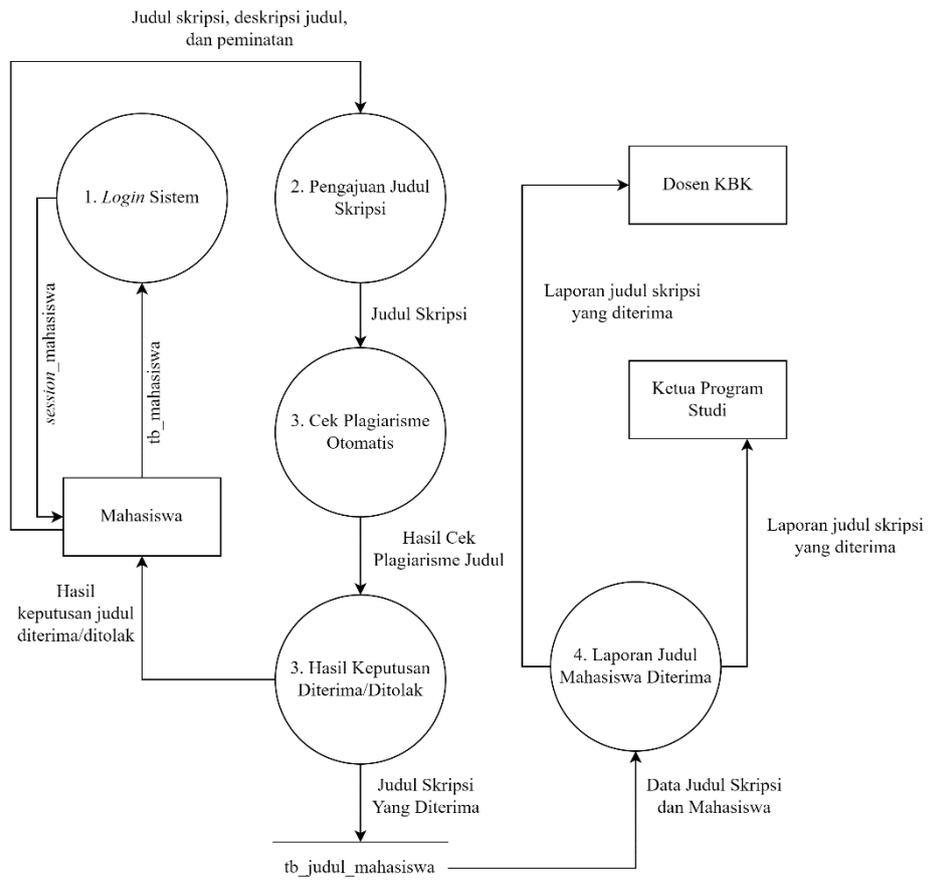
Diagram konteks merupakan diagram tingkatan paling atas yang menjelaskan hubungan sistem secara umum. Diagram konteks dari sistem pengajuan judul skripsi yang diajukan dapat dilihat pada gambar 3.4.



Gambar 3. 4 Diagram Konteks Sistem Yang Diajukan

3.1.2.2. Data Flow Diagram Sistem Yang Diajukan

Data flow diagram merupakan diagram yang menjelaskan lebih mendetail daripada diagram konteks. *Data Flow Diagram* level 1 dari sistem yang diajukan untuk pengajuan judul skripsi dapat dilihat pada gambar 3.5.



Gambar 3. 5 DFD Level 1 Sistem Yang Diajukan

3.2. Metode Pengembangan Perangkat Lunak

Pada penelitian ini menggunakan metode pengembangan perangkat lunak dengan model *prototype*. Metode ini sangat cocok untuk menyelesaikan penelitian ini karena penelitian ini hanya dilakukan dalam jangka waktu beberapa bulan saja (singkat).

3.2.1. Komunikasi

Tahapan ini mengumpulkan data sesuai dengan kebutuhan. Data yang dibutuhkan dalam penelitian ini adalah judul skripsi pada program studi Teknik Informatika Institut Informatika dan Bisnis Darmajaya dalam rentang tahun 3 tahun terakhir. Dalam tahap komunikasi ini peneliti juga menggunakan beberapa metode untuk mengumpulkan data, yaitu sebagai berikut.

3.2.1.1. Wawancara

Pada penelitian ini akan dilakukan wawancara dengan dosen Kelompok Bagian Keahlian (KBK) Institut Informatika dan Bisnis Darmajaya. Tujuan dari dilakukannya wawancara ini adalah untuk mengambil data mengenai persentase dari kemiripan judul skripsi yang akan dijadikan sebagai aturan dalam hasil akhir pada sistem pendeteksian tingkat kemiripan judul skripsi.

3.2.1.2. Studi Literatur

Pada penelitian ini peneliti mengambil serta menggunakan literatur untuk menunjang serta mendukung metode dengan cara mempelajari serta memahami berbagai teori dan konsep yang diambil dari buku, jurnal, *e-book*, skripsi, dan literatur lainnya yang berkaitan dengan penyelesaian masalah dalam penelitian ini.

3.2.2. Perencanaan Cepat

Berdasarkan hasil dari tahapan komunikasi maka selanjutnya adalah analisis kebutuhan untuk menunjang proses berjalannya penelitian ini. Adapun analisis kebutuhan yang dibutuhkan adalah sebagai berikut.

3.2.2.1. Analisis Kebutuhan Fungsional

Kebutuhan fungsional merupakan kebutuhan atas proses-proses yang akan diterapkan pada sistem agar sistem tersebut dapat berjalan dengan baik dan sesuai dengan kebutuhan. Proses-proses tersebut adalah sebagai berikut.

1. Admin

Admin dapat melakukan proses sebagai berikut.

- a. Masuk ke dalam sistem.
- b. Mengakses menu beranda.
- c. Tambah data mahasiswa.
- d. Tambah data judul skripsi.
- e. Tambah data pengguna.
- f. Mengelola pengajuan judul skripsi mahasiswa.
- g. Mengelola data mahasiswa.
- h. Mengelola data bank judul skripsi.
- i. Cetak laporan pengajuan judul.
- j. Mengatur tingkat persentase kemiripan.

2. Mahasiswa

Mahasiswa dapat melakukan proses sebagai berikut.

- a. Masuk ke dalam sistem.
- b. Mengakses menu beranda.
- c. Mengajukan judul.
- d. Melihat serta mengelola riwayat pengajuan judul.

3. Ketua Program Studi

Ketua Program Studi dapat melakukan proses sebagai berikut.

- a. Masuk kedalam sistem.
- b. Mengakses menu beranda.
- c. Mengelola pengajuan judul skripsi mahasiswa.
- d. Cetak laporan pengajuan judul.

3.2.2.2. Analisis Kebutuhan Non-Fungsional

Kebutuhan non fungsional merupakan kebutuhan sistem yang akan menggambarkan properti perilaku sistem, seperti kebutuhan perangkat lunak (*software*) dan perangkat keras (*hardware*). Kebutuhan atas perangkat lunak dan perangkat keras adalah sebagai berikut.

1. Analisis Kebutuhan Perangkat Lunak

Perangkat lunak atau *software* yang digunakan untuk membuat serta menjalankan program pada penelitian ini adalah sebagai berikut.

- a. Windows 10 64-bit.
- b. PHP.
- c. HTML.
- d. MySQL.
- e. Visual Studio Code.

2. Analisis Kebutuhan Perangkat Keras

Perangkat keras atau *hardware* yang digunakan untuk membuat serta menjalankan program pada penelitian ini adalah sebagai berikut.

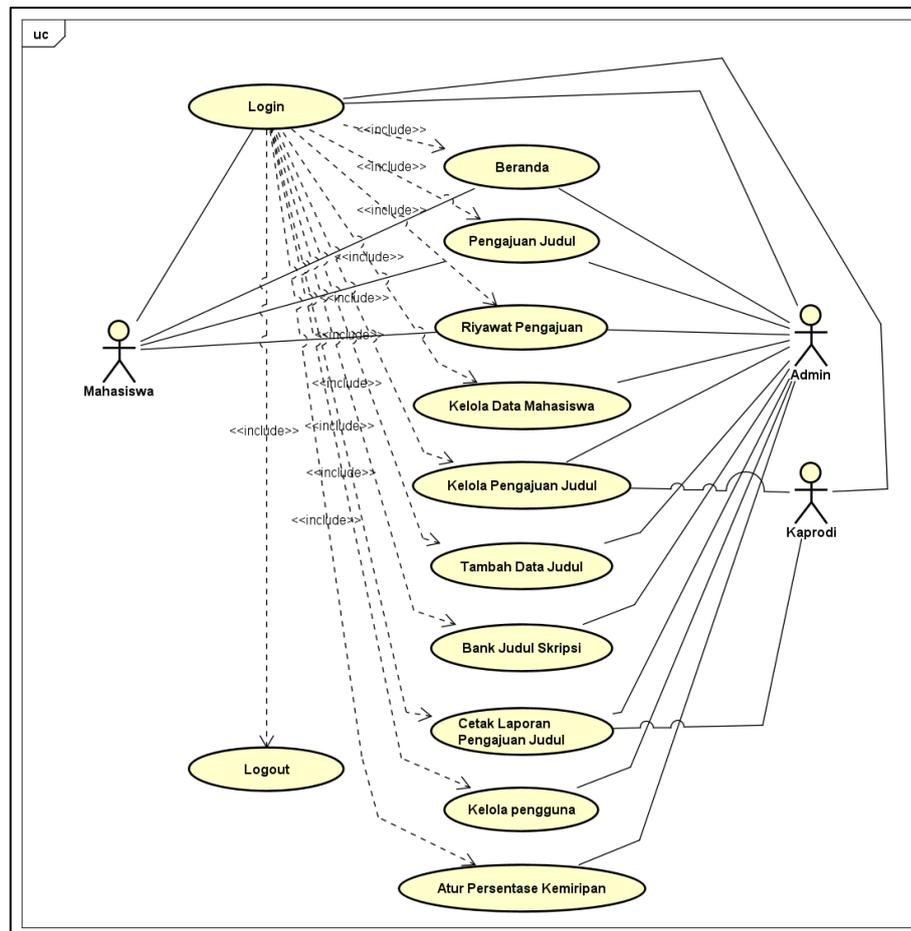
- a. Intel Core i3-4170 3.7 GHz
- b. NVIDIA GeForce GT 730 GDDR5 2GB VRAM.
- c. 8 GB DDR3 Memory 12800.
- d. 128 GB SSD & 512 GB HDD.
- e. Monitor Lenovo ThinkVision 23”.
- f. Keyboard dan Mouse.

3.2.3. Pemodelan Perancangan Cepat

Berdasarkan hasil dari tahapan *quick plan* maka selanjutnya adalah melakukan pemodelan. Pada tahap ini yang akan dilakukan adalah bagaimana cara kerja program, serta rancangan aplikasi yang diperoleh dari tahapan *quick plan* yang telah dianalisis dengan menggunakan *unified modified language* (UML), yaitu *use case diagram*, *activity diagram*, dan *class diagram* adalah sebagai berikut.

3.2.3.1. Use Case Diagram

Pada pembangunan sistem pendeteksi kemiripan judul skripsi berbasis *mobile*, dirancang sebuah *use case diagram* yang akan menggambarkan fungsi-fungsi pada masing-masing peran. Dalam pembangunan sistem ini terdapat dua level pengguna yang berbeda, yaitu admin dan mahasiswa. Terdapat 8 interaksi yang dirancang, yaitu masuk, menambah data mahasiswa, menambah data judul skripsi, melihat pengajuan judul, mengelola data mahasiswa, melihat bank judul skripsi, mengecek judul skripsi untuk diajukan, dan mengelola riwayat pengajuan judul. Rancangan *use case diagram* dapat dilihat pada gambar 3.6.



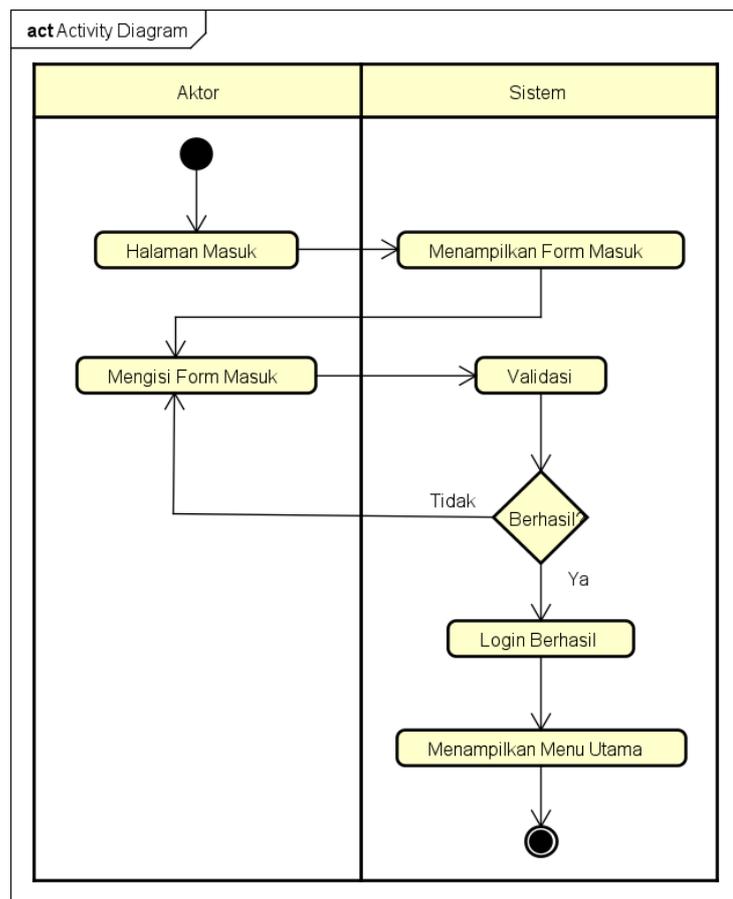
Gambar 3. 6 Use Case Diagram

3.2.3.2. Activity Diagram

Activity diagram berfungsi untuk menunjukkan setiap aktivitas yang terjadi pada sistem, baik aktivitas yang dapat dilakukan oleh admin maupun *user* (mahasiswa). Berikut adalah rancangan *activity diagram* untuk kedua aktor tersebut.

A. Activity Diagram Login

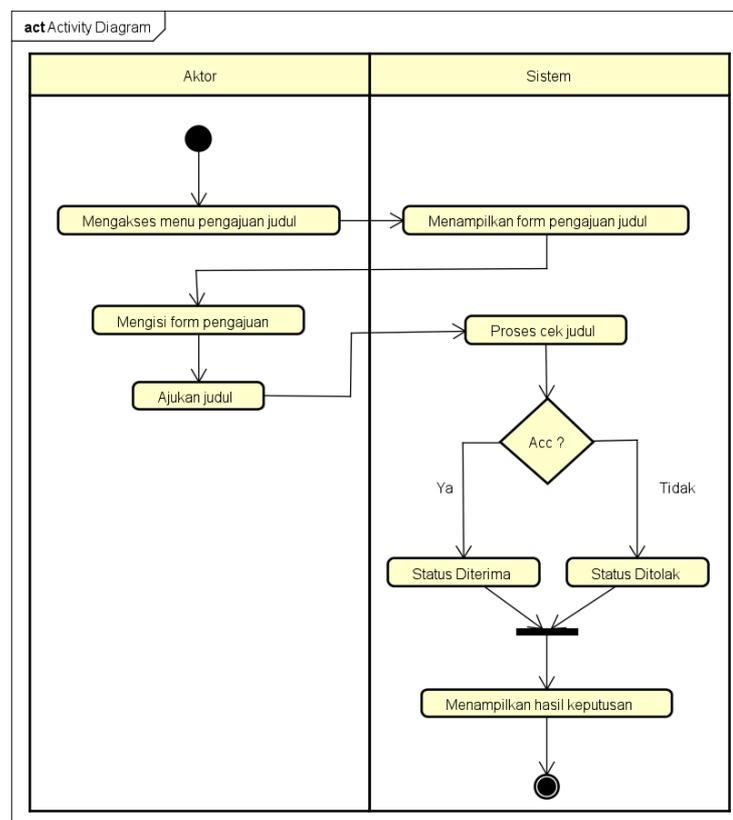
Activity Diagram Login menjelaskan bagaimana proses dalam melakukan validasi, jika akun *valid* maka sistem akan menampilkan halaman utama. Pada proses ini akan digunakan oleh admin dan mahasiswa untuk melakukan aktivitas *login* akun agar dapat menggunakan sistem. *Activity Diagram Login* dapat dilihat pada gambar 3.7.



Gambar 3. 7 Activity Diagram Login

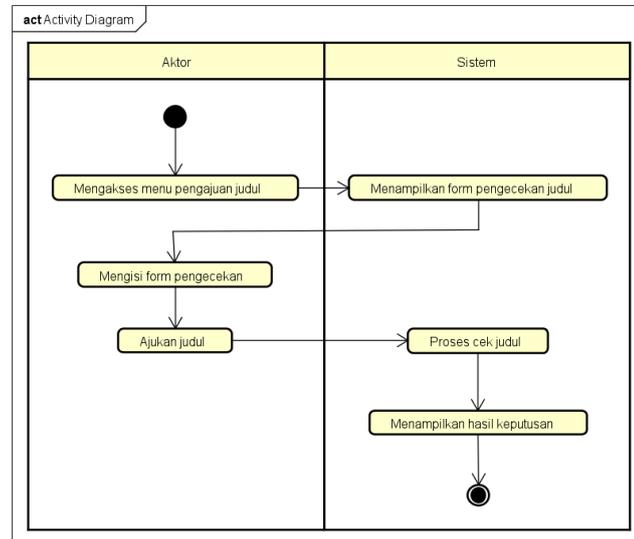
B. *Activity Diagram* Mengajukan dan Cek Judul

Activity Diagram Mengajukan dan Cek Judul menjelaskan bagaimana proses dalam mengajukan judul. Pada proses ini akan digunakan oleh mahasiswa terhadap judul yang diajukan akan dilakukan pengecekan oleh sistem dan hasil akhirnya akan diberi keputusan apakah judul dapat diajukan atau tidak. Sedangkan pada menu cek judul maka akan dilakukan pengecekan judul saja dan judul yang dilakukan pengecekan tidak diajukan dan tidak disimpan kedalam *database* sistem. *Activity Diagram* Mengajukan Judul dapat dilihat pada gambar 3.8.



Gambar 3. 8 *Activity Diagram* Pengajuan Judul

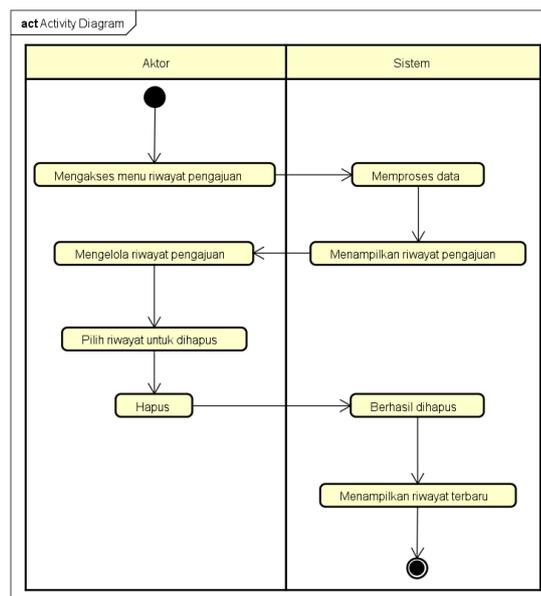
Activity Diagram Cek Judul dapat dilihat pada gambar 3.9.



Gambar 3. 9 Activity Diagram Cek Judul

C. Activity Diagram Riwayat Pengajuan

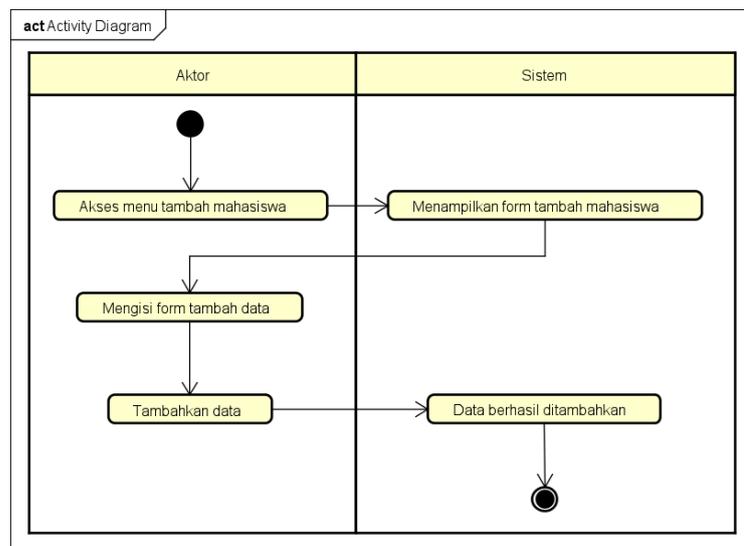
Activity Diagram Riwayat Pengajuan menjelaskan proses dalam melihat riwayat judul-judul yang sebelumnya diajukan dan pada proses ini akan digunakan oleh mahasiswa. Activity Diagram Riwayat Pengajuan dapat dilihat pada gambar 3.10.



Gambar 3. 10 Activity Diagram Riwayat Pengajuan

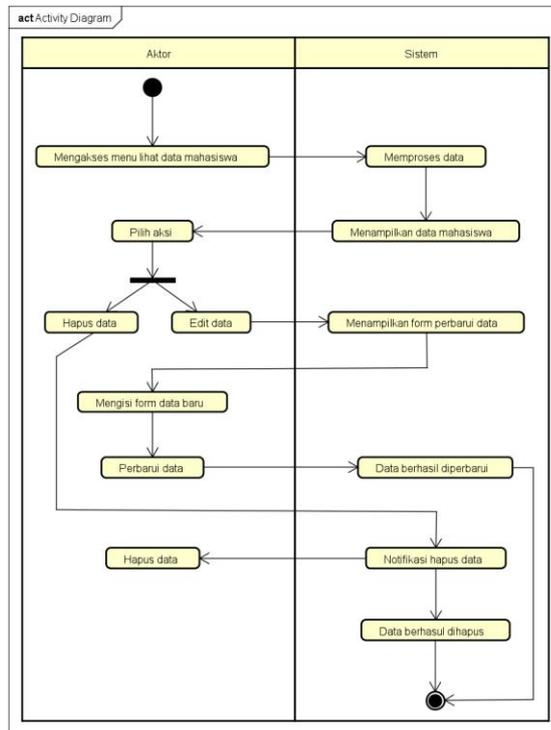
D. *Activity Diagram* Kelola Data Mahasiswa

Activity Diagram Kelola Data Mahasiswa menjelaskan proses dalam mengelola data mahasiswa. Pada proses ini akan digunakan oleh admin. Terdapat 3 proses yang dapat dilakukan, yaitu lihat data mahasiswa, edit data mahasiswa, dan hapus data mahasiswa. *Activity Diagram* Tambah Data Mahasiswa dapat dilihat pada gambar 3.11.



Gambar 3. 11 *Activity Diagram* Tambah Data Mahasiswa

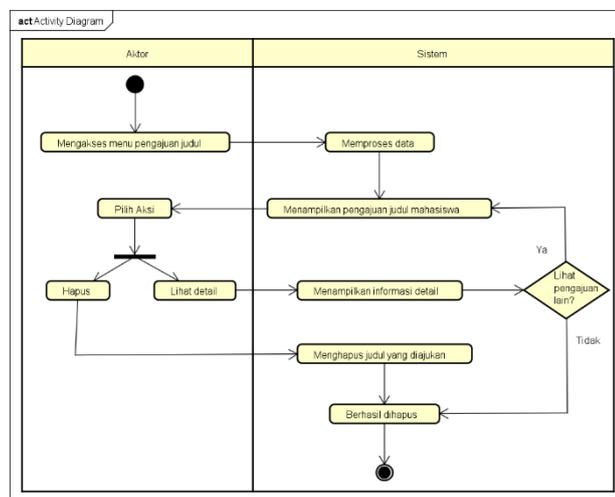
Activity Diagram Ubah dan Hapus Data Mahasiswa dapat dilihat pada gambar 3.12.



Gambar 3. 12 *Activity Diagram* Kelola Data Mahasiswa

E. *Activity Diagram* Lihat Pengajuan Judul Mahasiswa

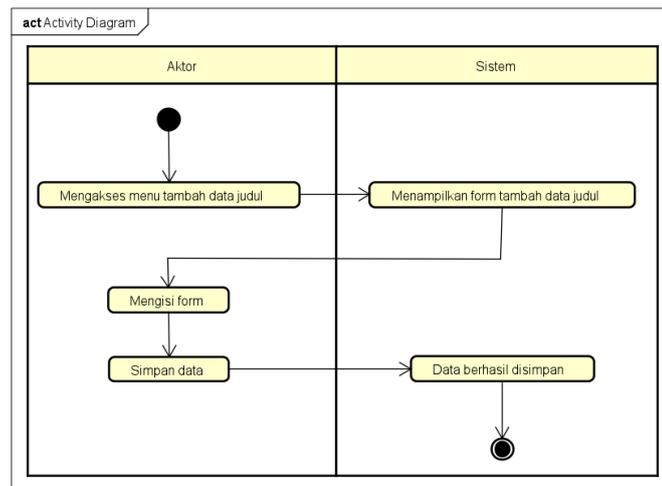
Activity Diagram Lihat Pengajuan Judul Mahasiswa menjelaskan proses untuk melihat judul-judul yang diajukan oleh para mahasiswa. Proses ini akan digunakan oleh admin. *Activity Diagram* Lihat Pengajuan Judul Mahasiswa dapat dilihat pada gambar 3.13.



Gambar 3. 13 *Activity Diagram* Pengajuan Judul Mahasiswa

F. *Activity Diagram* Tambah Data Judul

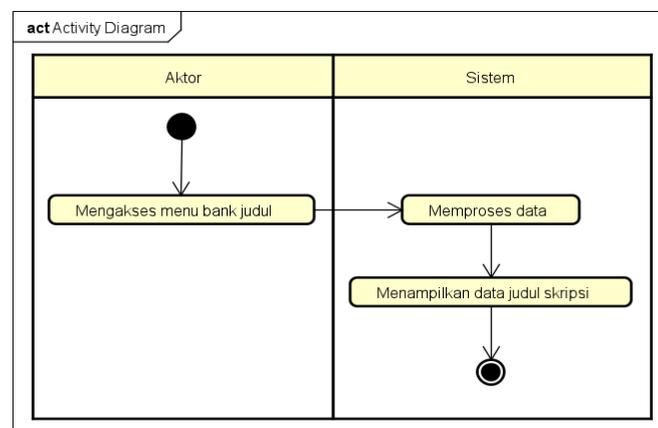
Activity Diagram Tambah Data Judul menjelaskan proses dalam menambahkan data judul skripsi kedalam database sistem dan pada proses ini akan digunakan oleh admin. *Activity Diagram* Tambah Data Judul dapat dilihat pada gambar 3.14.



Gambar 3. 14 *Activity Diagram* Tambah Data Judul

G. *Activity Diagram* Bank Judul Skripsi

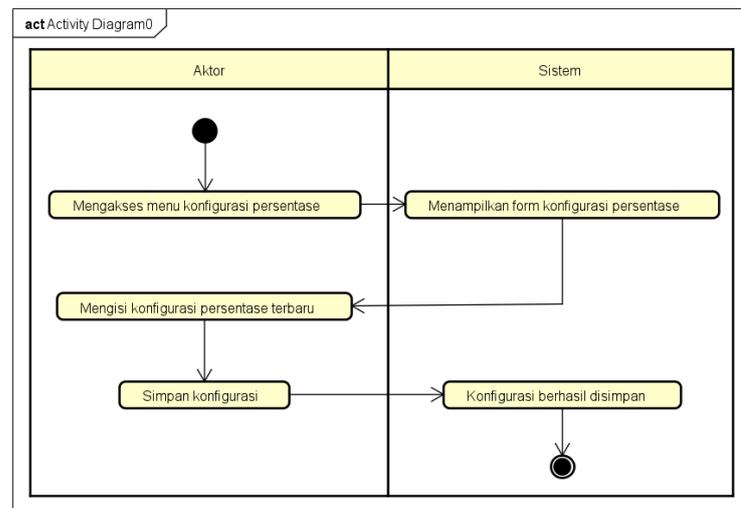
Activity Diagram Bank Judul Skripsi menjelaskan proses dalam melihat data-data judul skripsi yang ada pada database sistem. Pada proses ini akan digunakan oleh admin. *Activity Diagram* Bank Judul Skripsi dapat dilihat pada gambar 3.15.



Gambar 3. 15 *Activity Diagram* Bank Judul Skripsi

H. *Activity Diagram* Konfigurasi Persentase Kemiripan

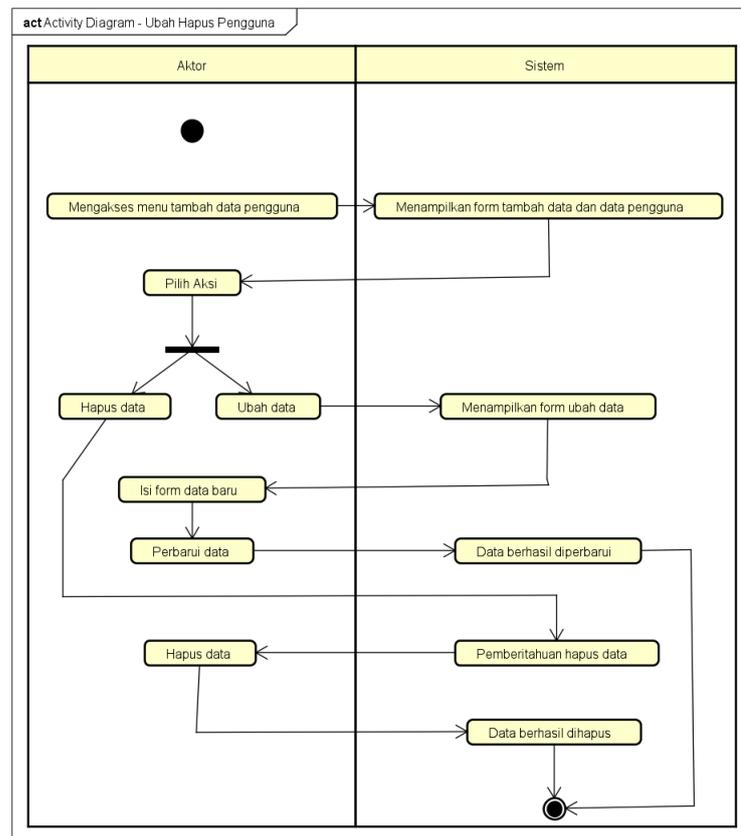
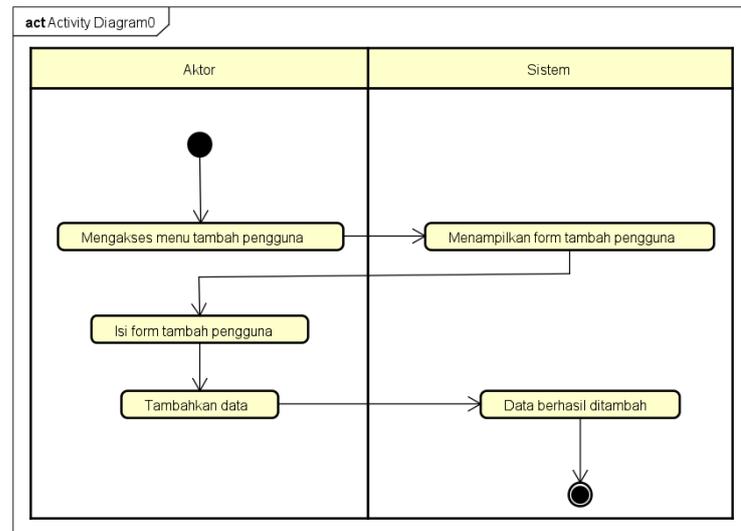
Activity Diagram Konfigurasi Persentase Kemiripan menjelaskan proses dalam melakukan perubahan dalam tingkat persentase kemiripan yang ditoleransi pada judul yang diajukan oleh mahasiswa. Proses ini akan digunakan oleh admin. *Activity Diagram* Konfigurasi Persentase Kemiripan dapat dilihat pada gambar 3.16.



Gambar 3. 16 *Activity Diagram* Konfigurasi Persentase Kemiripan

I. *Activity Diagram* Kelola Data Pengguna

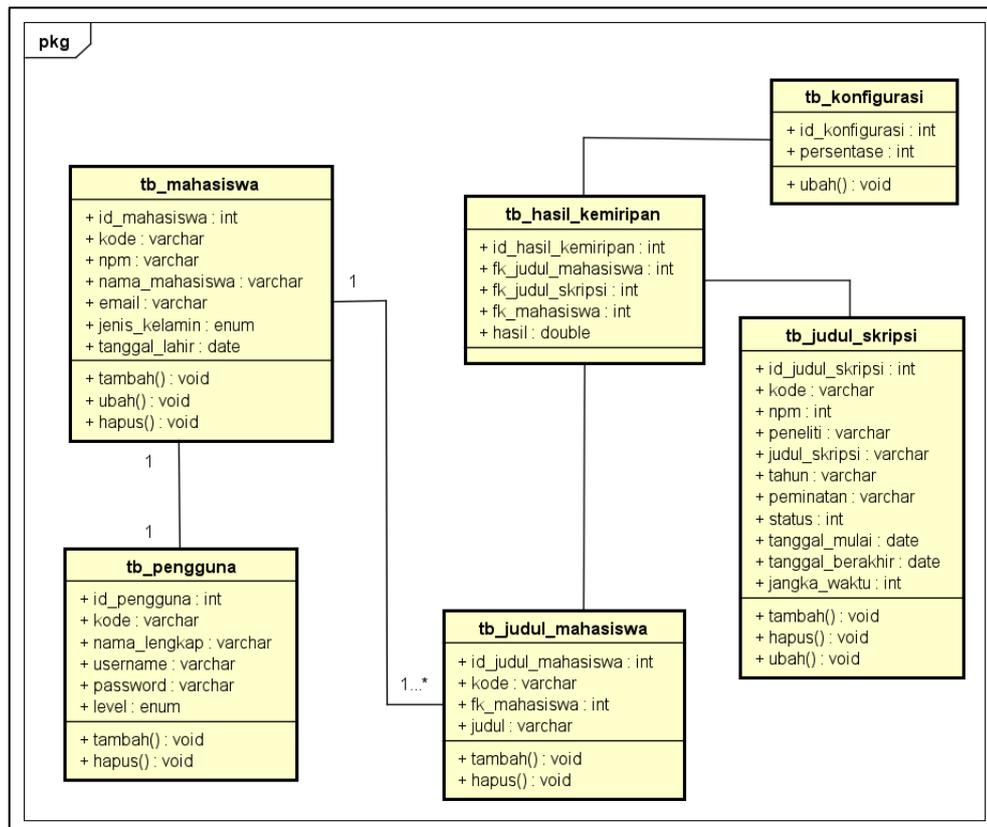
Activity Diagram Kelola Data Pengguna menjelaskan proses dalam melakukan pengelolaan data pengguna. Data pengguna yang dimaksud disini merupakan data pengguna untuk Ketua Program Studi, dan Dosen Kelompok Bagian Keilmuan (KBK). *Activity Diagram* Kelola Data Pengguna dapat dilihat pada gambar 3.17.



Gambar 3. 17 Activity Diagram Kelola Data Pengguna

3.2.3.3. Class Diagram

Class Diagram berfungsi untuk menggambarkan keadaan atribut atau properti dari suatu sistem. Rancangan *Class Diagram* dapat dilihat pada gambar 3.18.



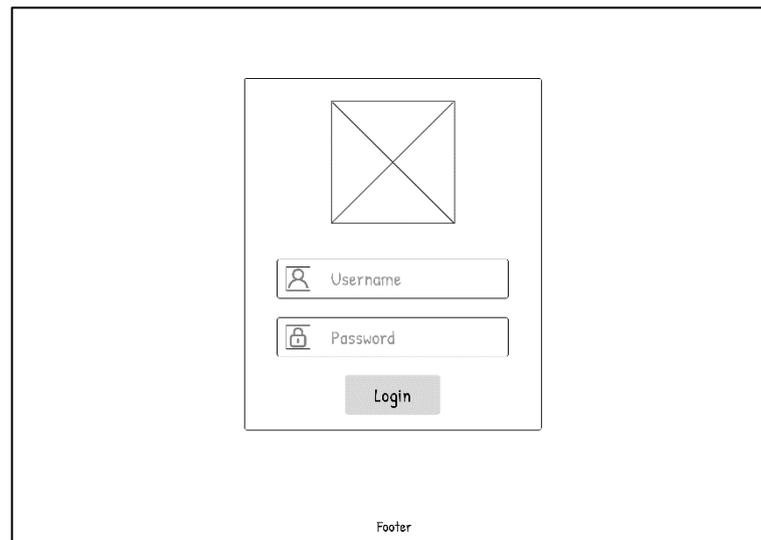
Gambar 3. 18 Class Diagram

3.2.4. Pembentukan Rancangan Prototipe

Berdasarkan hasil dari tahap *modelling* maka yang selanjutnya adalah tahapan *construction of prototype*. Pada tahapan ini merupakan proses implementasi dalam penyelesaian masalah pada penelitian ini yang akan dibuat berdasarkan tahapan-tahapan sebelumnya. Hasil dari analisis kebutuhan serta rancangan yang telah dilakukan kemudian akan diterapkan ke dalam bentuk program. Tampilan rancangan antarmuka dari sistem yang akan dibangun adalah sebagai berikut.

3.2.4.1. Halaman Masuk

Halaman masuk merupakan halaman yang berfungsi untuk melakukan verifikasi pengguna agar dapat masuk ke dalam *website*. Pengguna akan diminta untuk memasukkan nama pengguna dan kata sandi. Tampilan rancangan antarmuka dari halaman masuk dapat dilihat pada gambar 3.19.



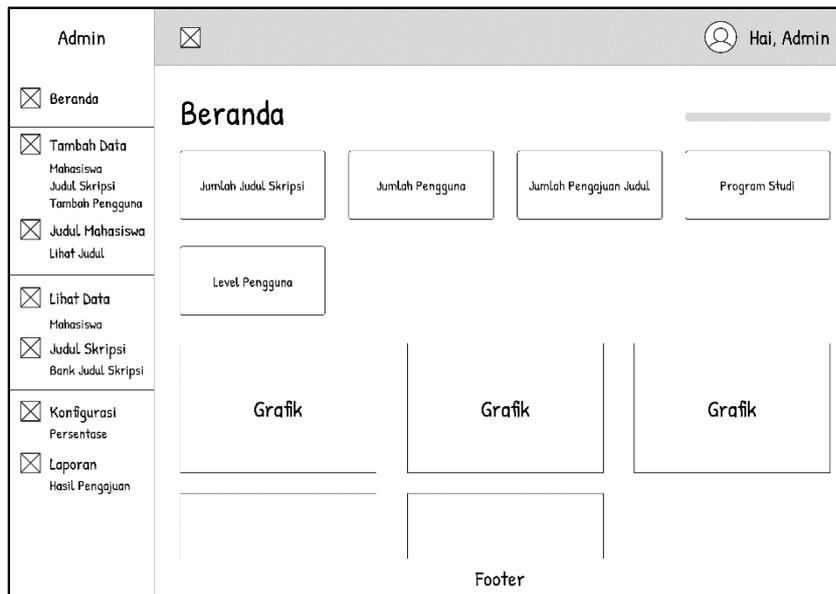
Gambar 3. 19 Rancangan Antarmuka Halaman Masuk

3.2.4.2. Halaman Hak Akses Admin

Halaman hak akses admin merupakan halaman-halaman yang dapat diakses oleh admin. Halaman-halaman yang dapat diakses oleh admin adalah sebagai berikut.

A. Halaman Beranda Admin

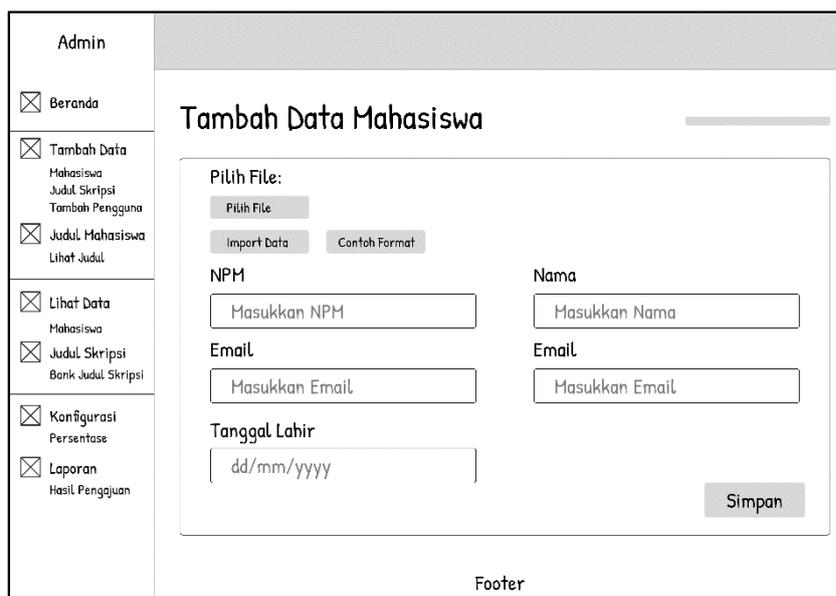
Halaman beranda admin berisi informasi jumlah judul skripsi, jumlah pengguna, jumlah judul yang diajukan oleh mahasiswa dan tampilan program studi. Pada menu ini juga terdapat grafik judul skripsi dari tahun ke tahun. Tampilan rancangan antarmuka dari halaman beranda admin dapat dilihat pada gambar 3.20.



Gambar 3. 20 Rancangan Antarmuka Beranda Admin

B. Halaman Tambah Data Mahasiswa

Halaman tambah data mahasiswa berfungsi untuk menambahkan data mahasiswa agar para mahasiswa dapat menggunakan website ini untuk melakukan pengecekan serta pengajuan judul skripsi. Tampilan rancangan antarmuka dari halaman tambah data mahasiswa dapat dilihat pada gambar 3.21.



Gambar 3. 21 Rancangan Antarmuka Tambah Data Mahasiswa

C. Halaman Tambah Data Judul Skripsi

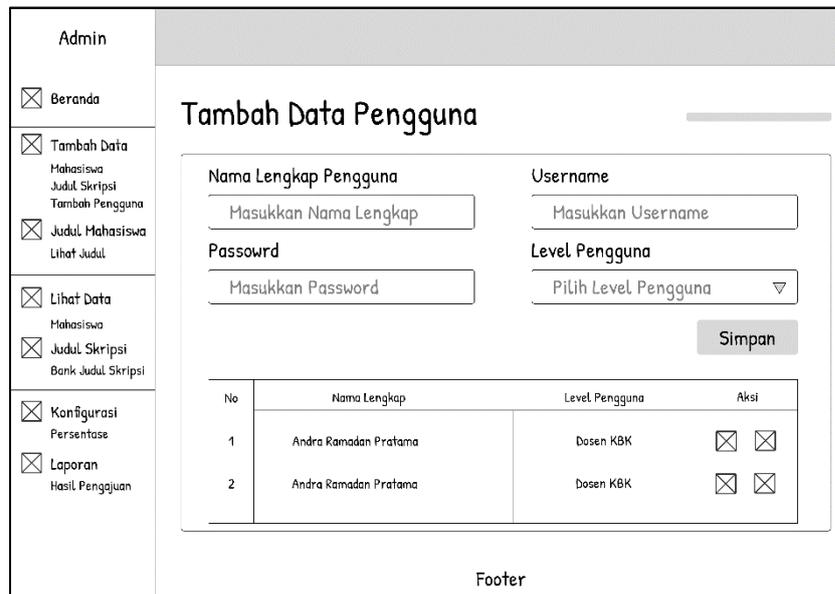
Halaman tambah data judul skripsi berfungsi untuk menambahkan data-data judul skripsi ke dalam *database website*. Tampilan rancangan antarmuka dari halaman tambah data judul skripsi dapat dilihat pada gambar 3.22.

The screenshot shows a web application interface for adding thesis titles. On the left is a sidebar menu with options like 'Beranda', 'Tambah Data Mahasiswa', 'Judul Skripsi', 'Judul Mahasiswa', 'Lihat Data Mahasiswa', 'Judul Skripsi', 'Bank Judul Skripsi', 'Konfigurasi Persentase', and 'Laporan Hasil Pengajuan'. The main content area is titled 'Tambah Data Judul' and contains a 'Pilih File:' section with 'Pilih File', 'Import Data', and 'Contoh Format' buttons. Below this are input fields for 'NPM' (Masukkan NPM), 'Peneliti' (Masukkan Peneliti), 'Judul' (Masukkan Judul), and 'Tahun' (Masukkan Tahun). At the bottom, there are dropdown menus for 'Peminatan' (Pilih Peminatan), 'Status' (Pilih Status), and 'Tanggal Mulai' (Tanggal Mulai), followed by a 'Simpan' (Save) button. A 'Footer' label is visible at the bottom center.

Gambar 3. 22 Rancangan Antarmuka Tambah Judul Skripsi

D. Halaman Tambah Data Pengguna

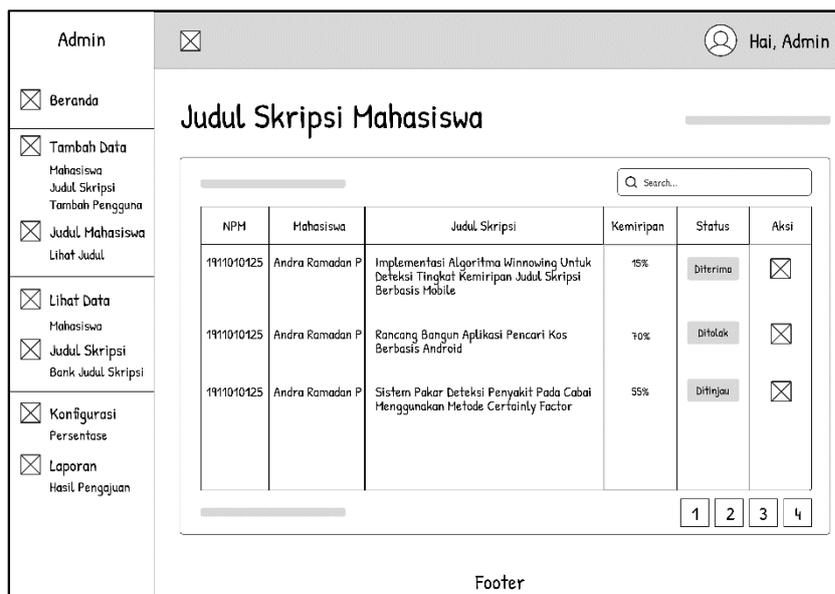
Halaman tambah data pengguna berfungsi untuk menambahkan data pengguna. Pengguna yang dimaksud adalah Ketua Program Studi dan Dosen Kelompok Bagian Keilmuan (KBK). Tampilan rancangan antarmuka dari halaman tambah data pengguna dapat dilihat pada gambar 3.23.



Gambar 3. 23 Rancangan Antarmuka Tambah Data Pengguna

E. Halaman Pengajuan Judul Mahasiswa

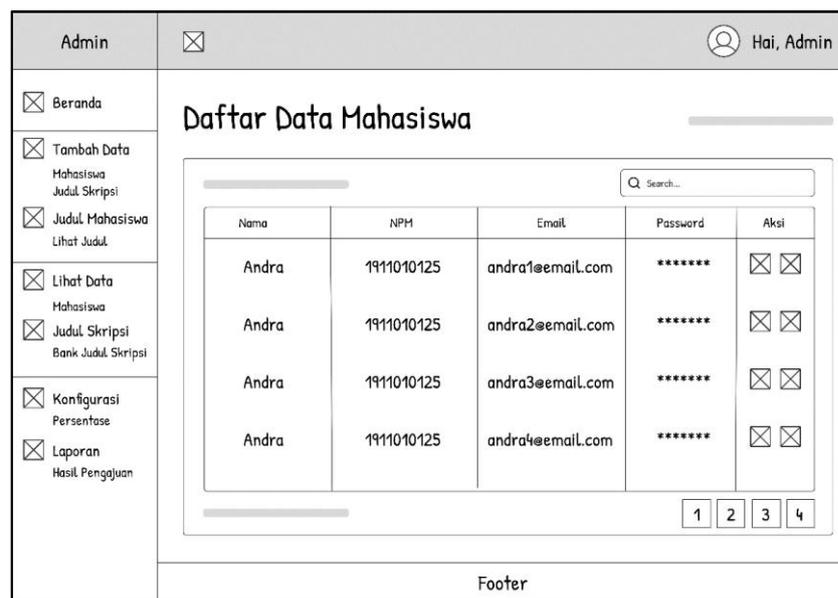
Halaman pengajuan judul mahasiswa merupakan halaman yang menampilkan judul-judul skripsi yang diajukan oleh mahasiswa. Tampilan rancangan antarmuka dari halaman pengajuan judul mahasiswa dapat dilihat pada gambar 3.24.



Gambar 3. 24 Rancangan Antarmuka Pengajuan Judul Mahasiswa

F. Halaman Lihat Data Mahasiswa

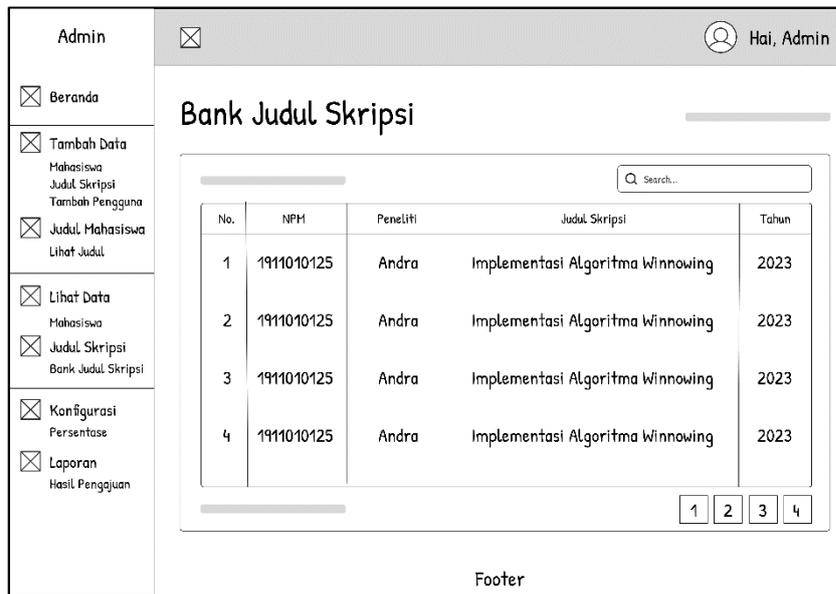
Halaman lihat data mahasiswa merupakan halaman yang menampilkan data-data dari mahasiswa sebagai pengguna website ini. Pada halaman ini admin dapat mencari, mengubah, serta menghapus data mahasiswa yang dipilih. Tampilan rancangan antarmuka dari halaman lihat data mahasiswa dapat dilihat pada gambar 3.25.



Gambar 3. 25 Rancangan Antarmuka Data Mahasiswa

G. Halaman Bank Judul Skripsi

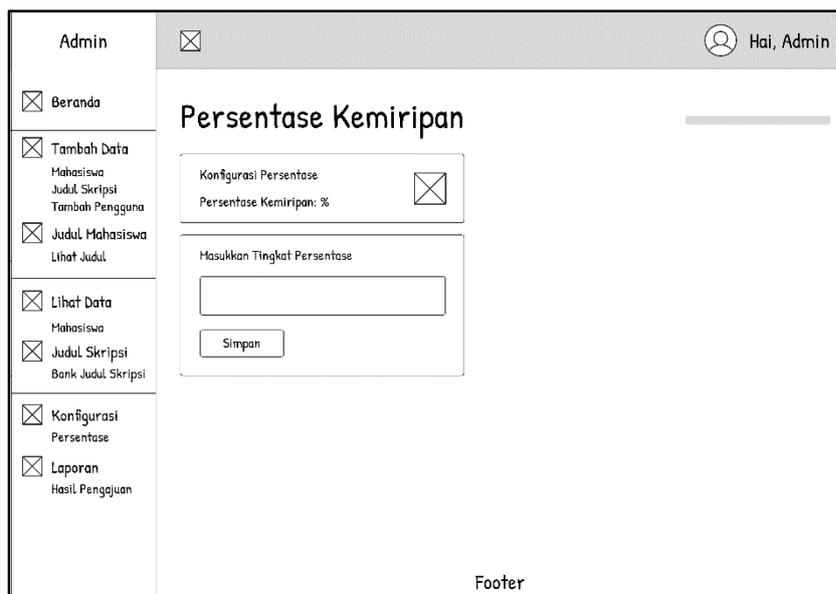
Halaman bank judul skripsi merupakan halaman dari data-data judul skripsi yang telah dimasukkan pada halaman tambah data judul skripsi. Tampilan rancangan antarmuka dari halaman bank judul skripsi dapat dilihat pada gambar 3.26.



Gambar 3. 26 Rancangan Antarmuka Bank Judul Skripsi

H. Halaman Konfigurasi Persentase Kemiripan

Halaman konfigurasi persentase kemiripan merupakan halaman yang berfungsi untuk mengatur besaran persentase kemiripan judul skripsi yang akan diajukan oleh mahasiswa. Tampilan rancangan antarmuka dari halaman konfigurasi persentase kemiripan dapat dilihat pada gambar 3.27.



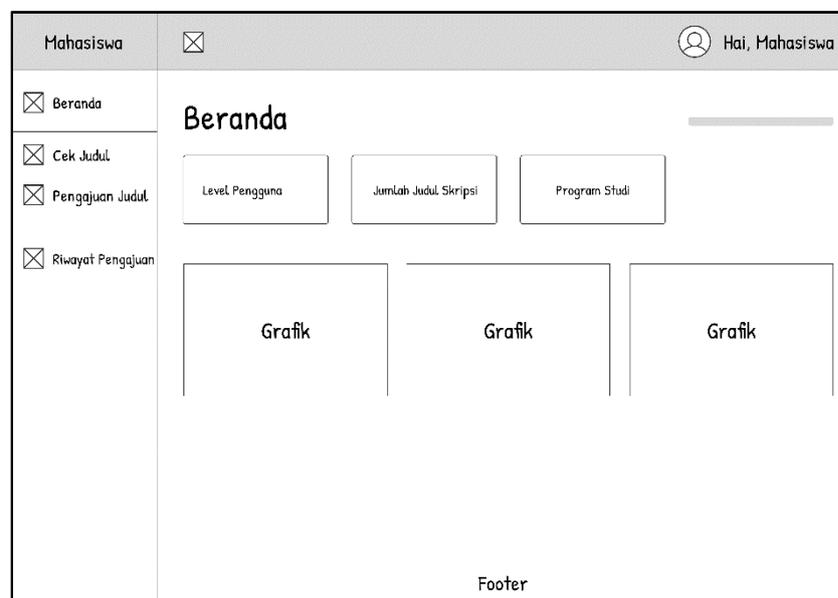
Gambar 3. 27 Rancangan Antarmuka Konfigurasi Persentase Kemiripan

3.2.4.3. Halaman Hak Akses Mahasiswa

Halaman hak akses mahasiswa merupakan halaman-halaman yang dapat diakses oleh mahasiswa. Halaman-halaman yang dapat diakses oleh admin adalah sebagai berikut.

A. Halaman Beranda Mahasiswa

Halaman beranda mahasiswa berisi informasi jumlah judul skripsi, dan tampilan program studi. Pada menu ini juga terdapat grafik judul skripsi dari tahun ke tahun. Tampilan rancangan antarmuka dari halaman beranda mahasiswa dapat dilihat pada gambar 3.28.



Gambar 3. 28 Rancangan Antarmuka Beranda Mahasiswa

B. Halaman Cek Judul

Halaman cek judul merupakan halaman yang berfungsi untuk melakukan pengecekan judul saja dan tidak mengajukan judul yang telah di lakukan pengecekan. Tampilan rancangan antarmuka dari halaman cek judul dapat dilihat pada gambar 3.29.



Gambar 3. 29 Rancangan Antarmuka Cek Judul

C. Halaman Pengajuan Judul

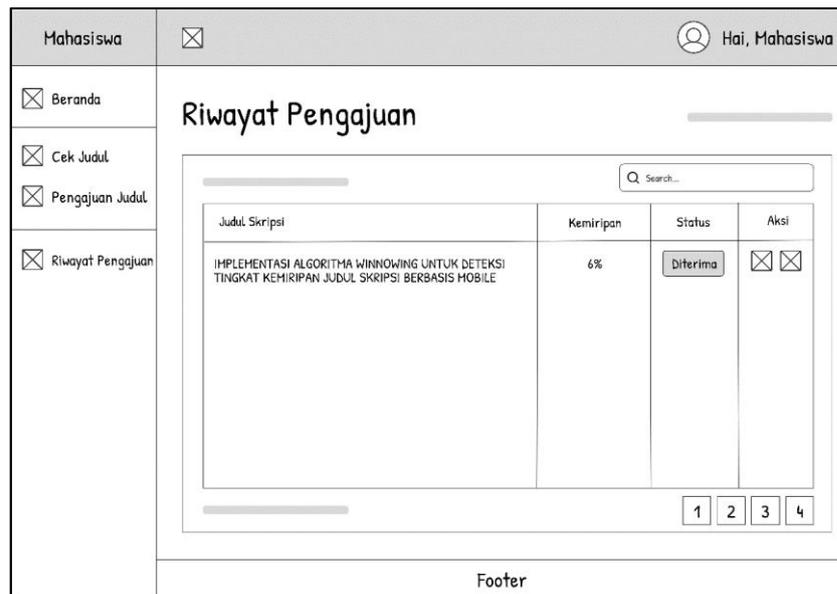
Halaman pengajuan judul merupakan halaman bagi mahasiswa untuk melakukan pengajuan judul skripsi. Pada halaman ini akan dilakukan pengecekan terhadap judul yang diajukan dan akan langsung menampilkan hasil apakah judul diterima atau ditolak. Tampilan rancangan antarmuka dari halaman pengajuan judul dapat dilihat pada gambar 3.30.



Gambar 3. 30 Rancangan Antarmuka Pengajuan Judul

D. Halaman Riwayat Pengajuan Judul

Halaman riwayat pengajuan judul merupakan halaman yang berisi riwayat judul-judul yang diajukan oleh mahasiswa. Tampilan rancangan antarmuka dari halaman riwayat pengajuan judul dapat dilihat pada gambar 3.31.



Gambar 3. 31 Rancangan Antarmuka Riwayat Pengajuan Judul

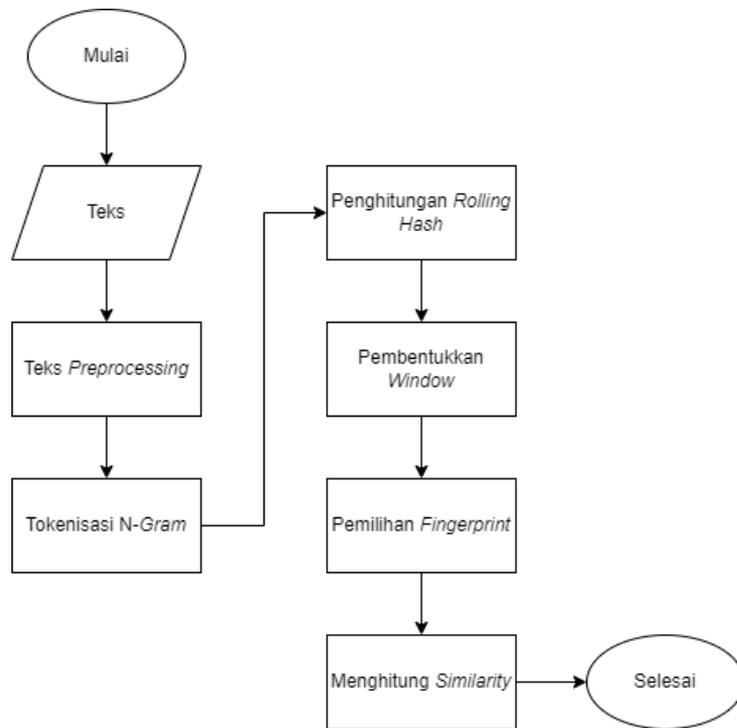
3.2.5. Penyerahan Sistem dan Umpan Balik

Tahapan ini merupakan tahap penyerahan serta pengujian perangkat lunak yang telah dibuat untuk memastikan kemampuan dari program untuk menyelesaikan masalah penelitian ini apakah sudah mencapai tujuan dan sudah sesuai dengan kebutuhan yang direncanakan.

3.3. Penerapan Algoritma *Winnower*

Pada algoritma *winnower* terdapat 6 tahapan, yaitu teks *preprocessing*, pembentukan *n-gram*, perhitungan *rolling hash*, pembentukan *window*, pemilihan *fingerprint* dari setiap *window*, dan yang terakhir adalah perhitungan *similarity*.

Untuk melakukan deteksi kemiripan antar teks atau dokumen dengan menggunakan algoritma *winnowing* maka akan dilakukan 6 tahapan. *Flowchart* dari tahapan penyelesaian menggunakan algoritma *winnowing* dapat dilihat pada gambar 3.32.



Gambar 3. 32 *Flowchart* Algoritma *Winnowing*

Sebagai contoh akan dilakukan proses deteksi kemiripan antara 2 judul skripsi. Judul pertama adalah “IMPLEMENTASI ALGORITMA *WINNOWER* UNTUK DETEKSI TINGKAT KEMIRIPAN JUDUL SKRIPSI BERBASIS *MOBILE*” dengan judul kedua “RANCANG BANGUN *WEBSITE* PROGRAM STUDI TEKNIK INFORMATIKA INSTITUT INFORMATIKA DAN BISNIS DARMAJAYA MENGGUNAKAN METODE *EXTREME PROGRAMING* (XP)”.

1. Teks *Preprocessing*

Pada tahapan ini akan dilakukan pembuangan karakter yang tidak relevan (*special character* dan *white space*). Maka kedua judul tersebut akan menjadi sebagai berikut.

Judul 1

implementasi algoritma windowing untuk deteksi tingkat kemiripan judul skripsi berbasis mobile

Judul 2

rancang bangun website program studi teknik informatika institut informatika dan bisnis darmajaya menggunakan metode extreme programming

2. Tokenisasi N-Gram

Tahapan selanjutnya adalah tokenisasi *n-gram*. Sebagai contoh maka akan digunakan besaran *n-gram* dengan $n = 5$. Maka tokenisasi *n-gram* dari kedua judul skripsi tersebut adalah sebagai berikut.

Judul 1

imple mplem pleme lemen ement menta entas ntasi tasia asial sialg ialgo
 algor lgori gorit oritm ritma itmaw tmawi mawin awinn winno innow nnowi
 nowin owinw wingu ingun ngunt guntu untuk ntukd tukde ukdet kdete detek
 eteks teksi eksit ksiti sitin iting tingk ingka ngkat gkatk katke atkem tkemi
 kemir emiri mirip iripa ripan ipanj panju anjud njudu judul uduls dulsk ulskr
 lskri skrip krips ripsi ipsib psibe siber iberb berba erbas rbasi basis asism
 sismo ismob smobi mobil obile

Judul 2

ranca ancan ncang cangb angba ngban gbang bangu angun ngunw gunwe
 unweb nwebs websi ebsit bsite sitep itepr tepro eprog progr rogra ogram
 grams ramst amstu mstud studi tudit udite ditekn tekni kniki niki
 ikinf kinfo infor nform forma ormat rmati matik atika tikai ikain kains ainst
 insti nstit stitu titut ituti tutin utinf tinfo infor nform forma ormat rmati matik
 atika tikad ikada kadan adanb danbi anbis nbisn bisni isnis snisd nisda isdar
 sdarm darma armaj rmaja majay ajaya jayam ayame yamen ameng mengg
 enggu nggun gguna gunakan unaka nakan akanm kanme anmet nmeto metod

etode todee odeex deext eextr extre xtrem treme remep emepr mepro eprog
 progr rogra ogram grami ramin aming mingx ingxp

3. Penghitungan *Rolling Hash*

Tahapan selanjutnya adalah merubah string-string yang telah terbentuk pada tahapan tokenisasi *n-gram* ke dalam bentuk desimal dengan menggunakan teknik *hashing*. Dengan menggunakan table ASCII untuk mengubah semua *string* karakter tersebut secara bergulir (*rolling hash*). Berikut adalah rumus yang digunakan untuk melakukan penghitungan *hashing*.

$$Hash = c_1 \times b^{(n-1)} + c_2 \times b^{(n-2)} + \dots + c_n \times b^{(n-n)} \times b$$

Keterangan:

c = Indeks karakter pada tabel ASCII

b = Bilangan prima

n = Besaran n-gram

Sebagai contoh untuk melakukan hashing pada string character “implementasi” dengan b = 2 dan n = 5, maka tokenisasi n-gram pertama adalah “imple” apabila dimasukkan kedalam rumus akan menjadi sebagai berikut.

$$Hash = (105 \times 2^{(5-1)} + 109 \times 2^{(4-1)} + 112 \times 2^{(3-1)} + 108 \times 2^{(2-1)} + 101 \times 2^{(1-1)}) \times 2$$

$$Hash = (105 \times 16 + 109 \times 8 + 112 \times 4 + 108 \times 2 + 101 \times 1) \times 2$$

$$Hash = (1680 + 872 + 448 + 216 + 101) \times 2$$

$$Hash = (3317) \times 2$$

$$Hash = 6634$$

Didapatkan hasil dari *string* karakter “imple” ke bentuk desimal dengan menggunakan teknik *hashing*, yaitu 6634. Proses ini dilakukan terus menerus hingga seluruh karakter dikonversi ke dalam bentuk desimal. Pada judul 1 dan

judul 2 dengan menggunakan besaran $n\text{-gram} = 5$ dan bilangan prima $b = 2$ akan menjadi seperti sebagai berikut.

Judul 1

6634 6766 6758 6568 6456 6642 6538 6822 6798 6388 6774 6410 6328 6658
 6636 6898 6886 6714 6918 6632 6508 7030 6682 6854 6888 6942 7014 6632
 6776 6746 7114 6940 7042 6892 6498 6362 6554 6854 6516 6778 6928 6702
 6898 6566 6644 6462 6534 6438 6878 6560 6482 6724 6666 6832 6580 6674
 6380 6786 6748 6942 6610 7048 6818 6948 6766 6894 6688 6858 6776 6388
 6250 6458 6662 6258 6462 6938 6712 6914 6684 6594

Judul 2

6670 6264 6526 6208 6274 6560 6286 6214 6376 6782 6726 7056 6854 6878
 6372 6482 6916 6700 6902 6586 6936 6898 6718 6562 6764 6466 6924 7082
 7036 6850 6426 6672 6834 6458 6662 6696 6556 6614 6608 6714 6582 6868
 6842 6602 6422 6846 6488 6486 6356 6714 6940 7074 7020 6826 7152 7084
 6902 6608 6714 6582 6868 6842 6602 6422 6836 6442 6384 6116 6234 6298
 6608 6386 6730 6940 6714 6616 6730 6294 6400 6786 6518 6254 6518 6454
 6920 6302 6602 6462 6680 6514 6650 6902 6536 6250 6494 6372 6758 6676
 6578 6894 6604 6336 6500 6738 7230 6982 6764 6460 6678 6586 6936 6898
 6718 6542 6712 6334 6700 6648

4. Pembentukan *Window*

Setelah dilakukan proses *hashing* terhadap semua karakter yang selanjutnya adalah membentuk *window-window* sejumlah W . Sebagai contoh besaran *window* adalah 4, maka *window* pertama dari hasil proses teknik *hashing* adalah [6634,6766,6758,6568] untuk *window* selanjutnya mulai dari hasil proses teknik *hashing* ke dua, yaitu [6766,6758,6568,6456]. Maka *window-window* yang terbentuk dari hasil proses teknik *hashing* pada judul 1 dan judul 2 adalah sebagai berikut.

Judul 1

| | |
|-----------------------------|-----------------------------|
| W-1: {6634 6766 6758 6568} | W-41: {6928 6702 6898 6566} |
| W-2: {6766 6758 6568 6456} | W-42: {6702 6898 6566 6644} |
| W-3: {6758 6568 6456 6642} | W-43: {6898 6566 6644 6462} |
| W-4: {6568 6456 6642 6538} | W-44: {6566 6644 6462 6534} |
| W-5: {6456 6642 6538 6822} | W-45: {6644 6462 6534 6438} |
| W-6: {6642 6538 6822 6798} | W-46: {6462 6534 6438 6878} |
| W-7: {6538 6822 6798 6388} | W-47: {6534 6438 6878 6560} |
| W-8: {6822 6798 6388 6774} | W-48: {6438 6878 6560 6482} |
| W-9: {6798 6388 6774 6410} | W-49: {6878 6560 6482 6724} |
| W-10: {6388 6774 6410 6328} | W-50: {6560 6482 6724 6666} |
| W-11: {6774 6410 6328 6658} | W-51: {6482 6724 6666 6832} |
| W-12: {6410 6328 6658 6636} | W-52: {6724 6666 6832 6580} |
| W-13: {6328 6658 6636 6898} | W-53: {6666 6832 6580 6674} |
| W-14: {6658 6636 6898 6886} | W-54: {6832 6580 6674 6380} |
| W-15: {6636 6898 6886 6714} | W-55: {6580 6674 6380 6786} |
| W-16: {6898 6886 6714 6918} | W-56: {6674 6380 6786 6748} |
| W-17: {6886 6714 6918 6632} | W-57: {6380 6786 6748 6942} |
| W-18: {6714 6918 6632 6508} | W-58: {6786 6748 6942 6610} |
| W-19: {6918 6632 6508 7030} | W-59: {6748 6942 6610 7048} |
| W-20: {6632 6508 7030 6682} | W-60: {6942 6610 7048 6818} |
| W-21: {6508 7030 6682 6854} | W-61: {6610 7048 6818 6948} |
| W-22: {7030 6682 6854 6888} | W-62: {7048 6818 6948 6766} |
| W-23: {6682 6854 6888 6942} | W-63: {6818 6948 6766 6894} |
| W-24: {6854 6888 6942 7014} | W-64: {6948 6766 6894 6688} |
| W-25: {6888 6942 7014 6632} | W-65: {6766 6894 6688 6858} |
| W-26: {6942 7014 6632 6776} | W-66: {6894 6688 6858 6776} |
| W-27: {7014 6632 6776 6746} | W-67: {6688 6858 6776 6388} |
| W-28: {6632 6776 6746 7114} | W-68: {6858 6776 6388 6250} |
| W-29: {6776 6746 7114 6940} | W-69: {6776 6388 6250 6458} |
| W-30: {6746 7114 6940 7042} | W-70: {6388 6250 6458 6662} |
| W-31: {7114 6940 7042 6892} | W-71: {6250 6458 6662 6258} |

| | |
|-----------------------------|-----------------------------|
| W-32: {6940 7042 6892 6498} | W-72: {6458 6662 6258 6462} |
| W-33: {7042 6892 6498 6362} | W-73: {6662 6258 6462 6938} |
| W-34: {6892 6498 6362 6554} | W-74: {6258 6462 6938 6712} |
| W-35: {6498 6362 6554 6854} | W-75: {6462 6938 6712 6914} |
| W-36: {6362 6554 6854 6516} | W-76: {6938 6712 6914 6684} |
| W-37: {6554 6854 6516 6778} | W-77: {6712 6914 6684 6594} |
| W-38: {6854 6516 6778 6928} | |
| W-39: {6516 6778 6928 6702} | |
| W-40: {6778 6928 6702 6898} | |

Judul 2

| | |
|-----------------------------|-----------------------------|
| W-1: {6670 6264 6526 6208} | W-59: {6714 6582 6868 6842} |
| W-2: {6264 6526 6208 6274} | W-60: {6582 6868 6842 6602} |
| W-3: {6526 6208 6274 6560} | W-61: {6868 6842 6602 6422} |
| W-4: {6208 6274 6560 6286} | W-62: {6842 6602 6422 6836} |
| W-5: {6274 6560 6286 6214} | W-63: {6602 6422 6836 6442} |
| W-6: {6560 6286 6214 6376} | W-64: {6422 6836 6442 6384} |
| W-7: {6286 6214 6376 6782} | W-65: {6836 6442 6384 6116} |
| W-8: {6214 6376 6782 6726} | W-66: {6442 6384 6116 6234} |
| W-9: {6376 6782 6726 7056} | W-67: {6384 6116 6234 6298} |
| W-10: {6782 6726 7056 6854} | W-68: {6116 6234 6298 6608} |
| W-11: {6726 7056 6854 6878} | W-69: {6234 6298 6608 6386} |
| W-12: {7056 6854 6878 6372} | W-70: {6298 6608 6386 6730} |
| W-13: {6854 6878 6372 6482} | W-71: {6608 6386 6730 6940} |
| W-14: {6878 6372 6482 6916} | W-72: {6386 6730 6940 6714} |
| W-15: {6372 6482 6916 6700} | W-73: {6730 6940 6714 6616} |
| W-16: {6482 6916 6700 6902} | W-74: {6940 6714 6616 6730} |
| W-17: {6916 6700 6902 6586} | W-75: {6714 6616 6730 6294} |
| W-18: {6700 6902 6586 6936} | W-76: {6616 6730 6294 6400} |
| W-19: {6902 6586 6936 6898} | W-77: {6730 6294 6400 6786} |
| W-20: {6586 6936 6898 6718} | W-78: {6294 6400 6786 6518} |

| | |
|-----------------------------|------------------------------|
| W-21: {6936 6898 6718 6562} | W-79: {6400 6786 6518 6254} |
| W-22: {6898 6718 6562 6764} | W-80: {6786 6518 6254 6518} |
| W-23: {6718 6562 6764 6466} | W-81: {6518 6254 6518 6454} |
| W-24: {6562 6764 6466 6924} | W-82: {6254 6518 6454 6920} |
| W-25: {6764 6466 6924 7082} | W-83: {6518 6454 6920 6302} |
| W-26: {6466 6924 7082 7036} | W-84: {6454 6920 6302 6602} |
| W-27: {6924 7082 7036 6850} | W-85: {6920 6302 6602 6462} |
| W-28: {7082 7036 6850 6426} | W-86: {6302 6602 6462 6680} |
| W-29: {7036 6850 6426 6672} | W-87: {6602 6462 6680 6514} |
| W-30: {6850 6426 6672 6834} | W-88: {6462 6680 6514 6650} |
| W-31: {6426 6672 6834 6458} | W-89: {6680 6514 6650 6902} |
| W-32: {6672 6834 6458 6662} | W-90: {6514 6650 6902 6536} |
| W-33: {6834 6458 6662 6696} | W-91: {6650 6902 6536 6250} |
| W-34: {6458 6662 6696 6556} | W-92: {6902 6536 6250 6494} |
| W-35: {6662 6696 6556 6614} | W-93: {6536 6250 6494 6372} |
| W-36: {6696 6556 6614 6608} | W-94: {6250 6494 6372 6758} |
| W-37: {6556 6614 6608 6714} | W-95: {6494 6372 6758 6676} |
| W-38: {6614 6608 6714 6582} | W-96: {6372 6758 6676 6578} |
| W-39: {6608 6714 6582 6868} | W-97: {6758 6676 6578 6894} |
| W-40: {6714 6582 6868 6842} | W-98: {6676 6578 6894 6604} |
| W-41: {6582 6868 6842 6602} | W-99: {6578 6894 6604 6336} |
| W-42: {6868 6842 6602 6422} | W-100: {6894 6604 6336 6500} |
| W-43: {6842 6602 6422 6846} | W-101: {6604 6336 6500 6738} |
| W-44: {6602 6422 6846 6488} | W-102: {6336 6500 6738 7230} |
| W-45: {6422 6846 6488 6486} | W-103: {6500 6738 7230 6982} |
| W-46: {6846 6488 6486 6356} | W-104: {6738 7230 6982 6764} |
| W-47: {6488 6486 6356 6714} | W-105: {7230 6982 6764 6460} |
| W-48: {6486 6356 6714 6940} | W-106: {6982 6764 6460 6678} |
| W-49: {6356 6714 6940 7074} | W-107: {6764 6460 6678 6586} |
| W-50: {6714 6940 7074 7020} | W-108: {6460 6678 6586 6936} |
| W-51: {6940 7074 7020 6826} | W-109: {6678 6586 6936 6898} |

| | |
|-----------------------------|------------------------------|
| W-52: {7074 7020 6826 7152} | W-110: {6586 6936 6898 6718} |
| W-53: {7020 6826 7152 7084} | W-111: {6936 6898 6718 6542} |
| W-54: {6826 7152 7084 6902} | W-112: {6898 6718 6542 6712} |
| W-55: {7152 7084 6902 6608} | W-113: {6718 6542 6712 6334} |
| W-56: {7084 6902 6608 6714} | W-114: {6542 6712 6334 6700} |
| W-57: {6902 6608 6714 6582} | W-115: {6712 6334 6700 6648} |
| W-58: {6608 6714 6582 6868} | |

5. Pemilihan *Fingerprint*

Setelah *window* terbentuk maka yang selanjutnya adalah proses pemilihan *fingerprint* yang nantinya akan dihitung kemiripannya antara judul 1 dengan judul 2. Pada proses pemilihan *fingerprint* ini akan dipilih nilai terkecil dari setiap *window*-nya. Sebagai contoh pada *window* pertama judul 1, yaitu [6634,6766,6758,6568] nilai terkecil dari *window* tersebut adalah 6568. Hal ini dilakukan terus menerus hingga seluruh *fingerprint* didapatkan dari *window-window* yang ada. Pemilihan *fingerprint* dari *window* yang terbentuk pada judul 1 dan judul 2 adalah sebagai berikut.

Judul 1

6568 6456 6456 6456 6456 6538 6388 6388 6388 6328 6328 6328 6328 6636
6636 6714 6632 6508 6508 6508 6508 6682 6682 6854 6632 6632 6632 6632
6746 6746 6892 6498 6362 6362 6362 6362 6516 6516 6516 6702 6566 6566
6462 6462 6438 6438 6438 6438 6482 6482 6482 6580 6580 6380 6380 6380
6380 6610 6610 6610 6610 6766 6766 6688 6688 6688 6388 6250 6250 6250
6250 6258 6258 6258 6462 6684 6594

Judul 2

6208 6208 6208 6208 6214 6214 6214 6214 6376 6726 6726 6372 6372 6372
6372 6482 6586 6586 6586 6586 6562 6562 6466 6466 6466 6466 6850 6426
6426 6426 6426 6458 6458 6458 6556 6556 6556 6582 6582 6582 6582 6422
6422 6422 6422 6356 6356 6356 6356 6714 6826 6826 6826 6826 6608 6608
6582 6582 6582 6582 6422 6422 6422 6384 6116 6116 6116 6116 6234 6298

| |
|---|
| 6386 6386 6616 6616 6294 6294 6294 6294 6254 6254 6254 6254 6302 6302 |
| 6302 6302 6462 6462 6514 6514 6250 6250 6250 6250 6372 6372 6578 6578 |
| 6336 6336 6336 6336 6500 6738 6460 6460 6460 6460 6586 6586 6542 6542 |
| 6334 6334 6334 |

6. Perhitungan *Similarity*

Pada algoritma *winnowing* tahap terakhir adalah menghitung nilai kemiripan dari teks atau dokumen yang dibandingkan. Dalam melakukan perhitungan menggunakan koefisien *jaccard*.

Diketahui:

- a. Jumlah *fingerprint* judul 1 = 77
- b. Jumlah *fingerprint* judul 2 = 115
- c. Gabungan (*union*) *fingerprint* judul 1 dan 2 = 192
- d. Irisan (*intersection*) *fingerprint* judul 1 dan 2 = 11
- e. Gabungan – Irisan (*union – intersection*) = 181

Maka perhitungan kemiripan menggunakan koefisien *jaccard* adalah sebagai berikut.

$$similarity = \frac{intersection}{(union - intersection)} \times 100$$

$$similarity = \frac{11}{(181)} \times 100$$

$$similarity = 6.08\%$$

Hasil yang didapatkan dari menggunakan algoritma *winnowing* untuk membandingkan judul 1 dengan judul 2 dengan kombinasi n-gram sebesar 5, bilangan prima sebesar 2 dan window sebesar 4 didapatkan hasil persentase kemiripan sebesar 6.08% dan dapat dinyatakan bahwa judul 1 dan judul 2 tidak sama.