

BAB III

ANALISIS DAN DESAIN SISTEM

III.1. Analisis Masalah

Pada tahap ini berisi tentang gambaran secara garis besar terhadap proses yang dilakukan aplikasi deteksi penjiplakan dokumen. Pada umumnya setiap algoritma pendeteksi penjiplakan memiliki tahapan-tahapan yang sama dalam pemrosesan dokumen teks. Tahapan yang dilakukan adalah adanya tahapan input, proses, dan output. Pada tahap input dilakukan pemasukan dokumen teks yang akan diuji. Selanjutnya dokumen akan diproses berdasarkan tahapan yang dimiliki oleh algoritma pendeteksi plagiarisme. Tahapan tersebut diantaranya adalah tahap preprocessing dan tokenisasi. Proses akan berlanjut dengan perhitungan tingkat similarity dokumen. Setelah proses utama dilakukan, selanjutnya aplikasi akan menghasilkan output informasi dokumen berupa hasil similarity dokumen dan kalimat yang telah diplagiasi.

III.2. Strategi Pemecahan Masalah

Sistem yang sedang berjalan saat ini masih terdapat beberapa kekurangan yang terdapat pada sistem yang telah ada sebelumnya :

1. Merancang dan membangun sebuah aplikasi yang dapat mempermudah pengguna dalam menentukan deteksi dokumen teks.
2. Menganalisa dan merencanakan sebuah aplikasi deteksi dokumen teks

dengan menggunakan android studio dan MYSQL yang dapat memberikan semua informasi yang kepada mahasiswa dan dosen.

III.3. Analisa Algoritma Winnowing

Pada tahap ini, analisa yang dilakukan adalah analisa terhadap algoritma winnowing asli yang belum dilakukan penambahan konsep biword. Prinsip kerja dari algoritma winnowing itu sendiri adalah menggunakan teknik hashing, yaitu mengkonversi setiap string menjadi bilangan dan kemudian bilangan tersebut akan disimpan sebagai fingerprint. Untuk memilih fingerprint dari hasil teknik hashing, dilakukan pembagian fingerprint ke dalam window tertentu dengan ukuran w .

Dari setiap window yang telah dibentuk, dipilih nilai hash yang paling minimum atau yang paling kecil. Jika terdapat nilai minimum lebih dari satu nilai, maka pilih dari window sebelah kanan. Kemudian simpan hasil hash yang telah dipilih yang merupakan fingerprint dokumen.

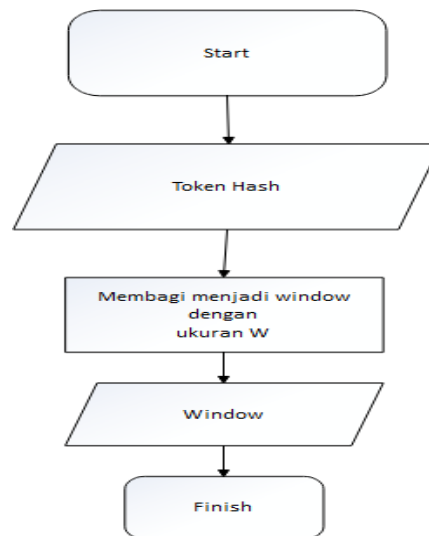
Secara umum prinsip kerja dari metode dokumen fingerprinting adalah sebagai berikut:

1. Hilangkan tanda baca dan spasi.
2. Sebelum melakukan fungsi hash, terlebih dahulu lakukan pembagian teks dokumen menjadi k -gram, dimana k merupakan parameter yang dipilih pengguna.
3. Menghitung nilai hash untuk setiap k -grams.
4. Memilih beberapa hasil hash menjadi dokumen fingerprinting

1. Membagi ke dalam beberapa *window*.

Token-token yang telah diperoleh, akan dibagi dalam beberapa *window* dengan ukuran w . Ukuran *window* ditentukan oleh pengguna aplikasi.

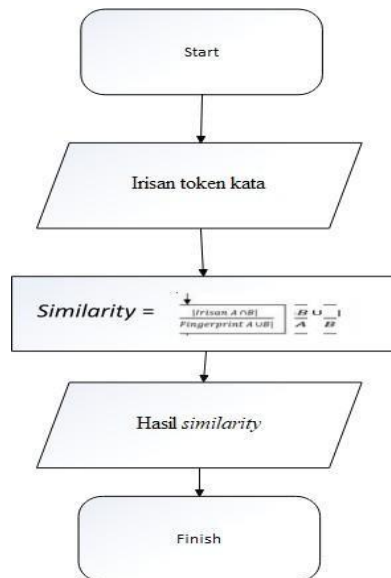
Berikut adalah *flowchart* pembentukan *window*:



Gambar 4.7 *Flowchart* proses pembentukan *window*

Nilai *fingerprint* yang diperoleh akan digunakan untuk menghitung *similarity* dokumen. Proses perhitungan dilakukan menggunakan persamaan *jaccard coefficient*.

Untuk lebih jelasnya dapat dilihat pada *flowchart* berikut:



Gambar 4.9 *Flowchart* proses hitung *similarity*

Contoh :

Misalkan terdapat ada dua kata “AKU” dan “ACU” untuk di bandingkan menggunakan algoritma winnowing, maka penyelesaian sebagai berikut :

1. Preprocessing

Aku | Acu

2. Rangkaian Gram

(b) – 3

Aku | Acu

3. Hash

$$A = 65 \times (3^3 - 1)$$

$$= 65 \times 3^2$$

$$= 65 \times 9$$

$$= 585$$

$$\begin{aligned} K &= 75 \times (3^3-2) = 75 \times 3 \\ &= 225 \end{aligned}$$

$$\begin{aligned} U &= 85 \times (3^3-3) = 85 \times 1 \\ &= 85 \end{aligned}$$

Setelah ini kita masuk di "ACU"

$$\begin{aligned} A &= 65 \times (3^3-1) \\ &= 65 \times 3^2 \\ &= 65 \times 9 \\ &= 585 \end{aligned}$$

$$\begin{aligned} C &= 67 \times (3^3-2) = 67 \times 3 \\ &= 201 \end{aligned}$$

$$\begin{aligned} U &= 85 \times (3^3-3) = 85 \times 1 \\ &= 85 \end{aligned}$$

4. Window

585 225 85 | 585 201 85

5. Fingerprint

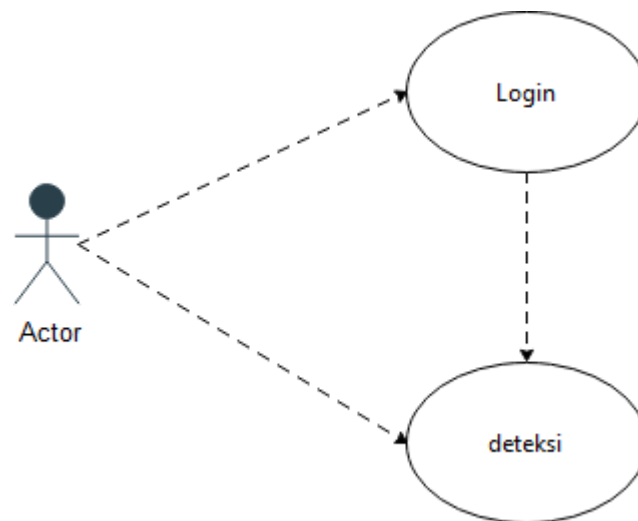
$$2/3 = 67 \%$$

III.4. Desain Sistem

Peneliti menggunakan pemodelan Unified Modelling Language (UML) yaitu Use Case, Class, Sequence dan Activity Diagram.

III.4.1. Use Case Diagram

Gambar 5.1 menjelaskan Use Case Diagram mengenai Penerapan aplikasi deteksi kemiripan dokumen teks untuk mendeteksi plagiasi.



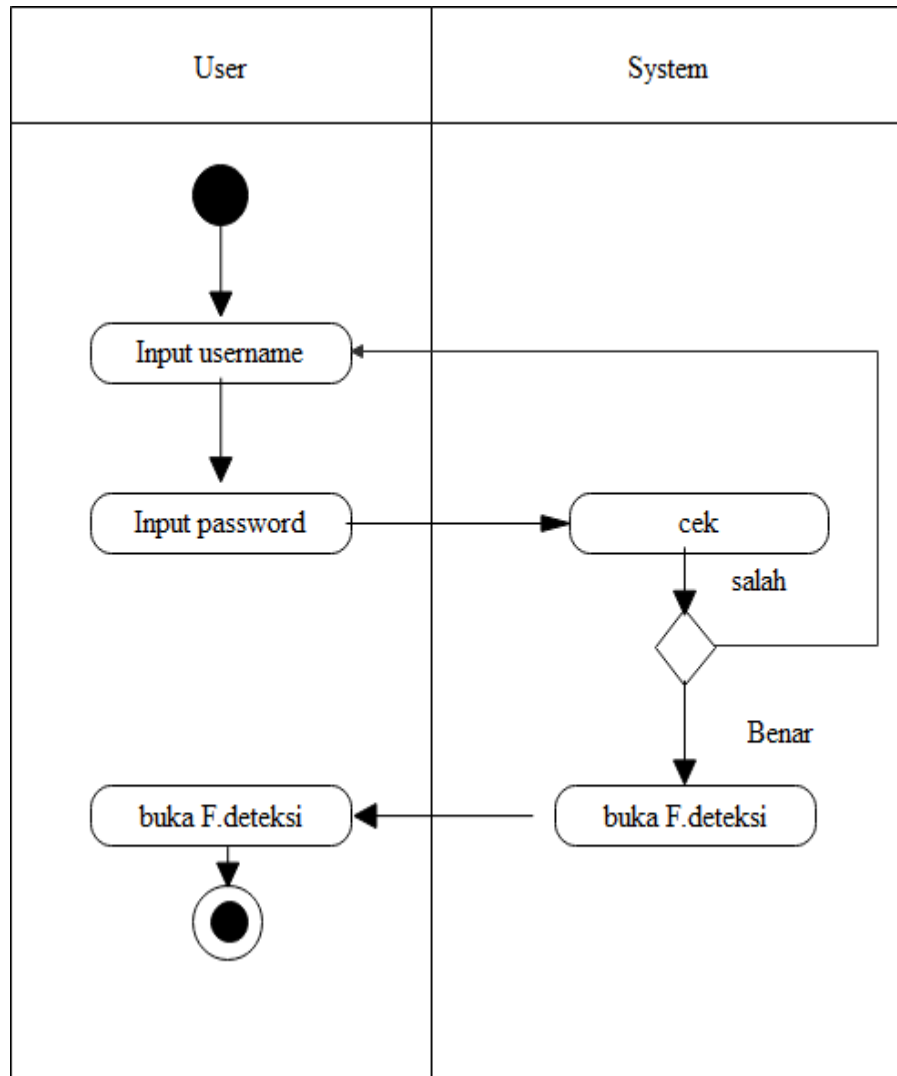
Gambar 5.1 Use Case Penerapan aplikasi deteksi kemiripan dokumen teks menggunakan algoritma winnowing

III.4.3. Activity Diagram

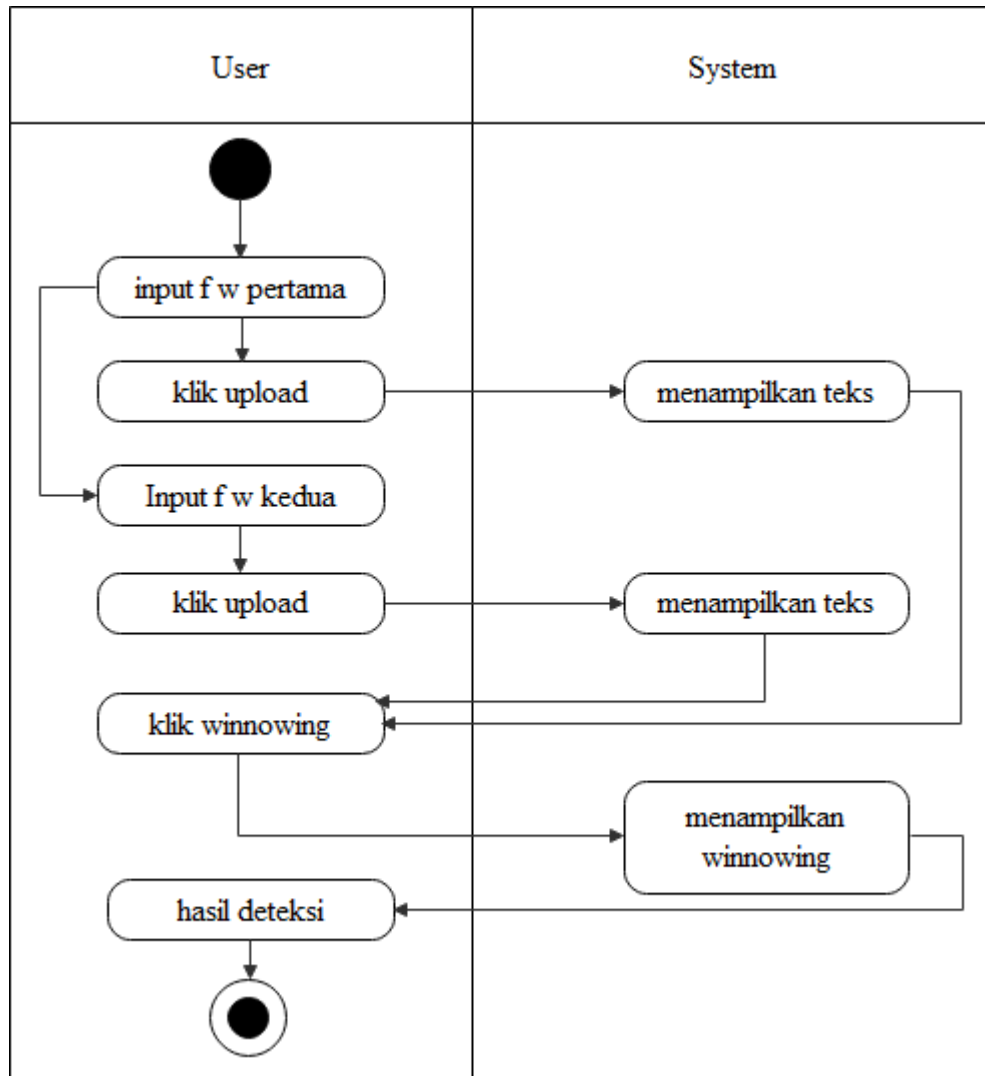
Activity Diagram Penerapan Penerapan aplikasi deteksi kemiripan dokumen teks menggunakan algoritma winnowing

2. Activity Diagram Input Data Kriteria

Activity diagram untuk menginput data kriteria dapat dilihat seperti pada gambar III.4 berikut :



Gambar 5.3. Activity Diagram login

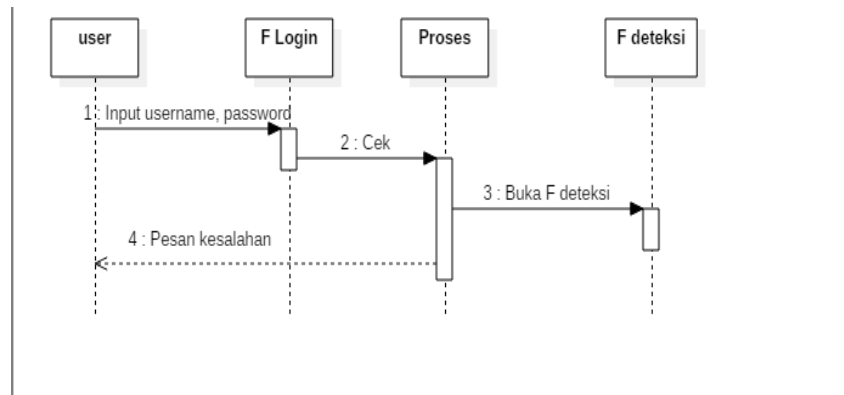


Gambar 5.4. Activity Diagram deteksi

III.4.4. Sequence Diagram

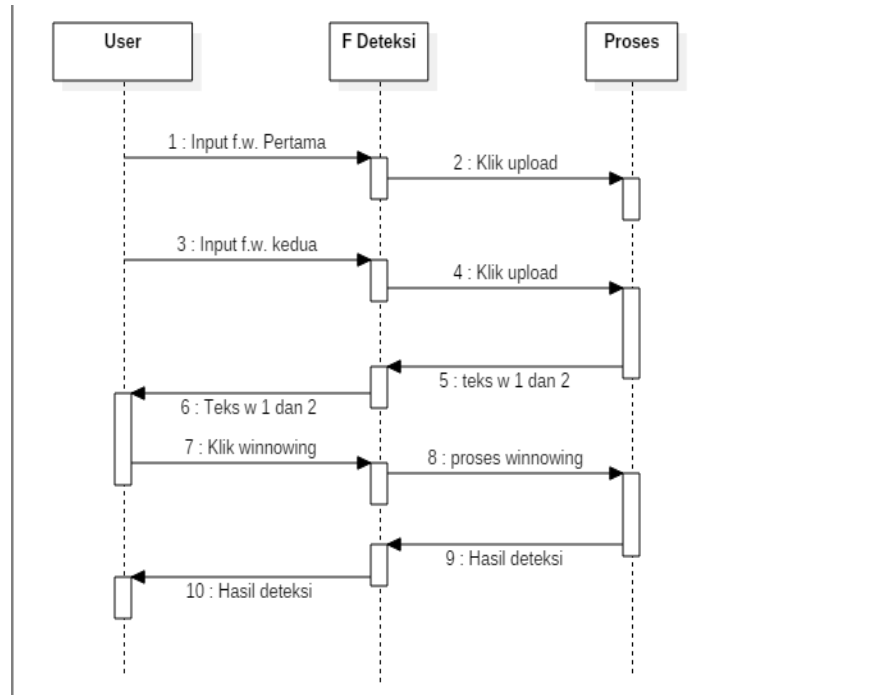
Sequence Diagram aplikasi deteksi kemiripan dokumen teks sebagai berikut:

2. Sequence Diagram Form login



Gambar 5.7. Sequence Diagram Form Login

3. Sequence Diagram deteksi

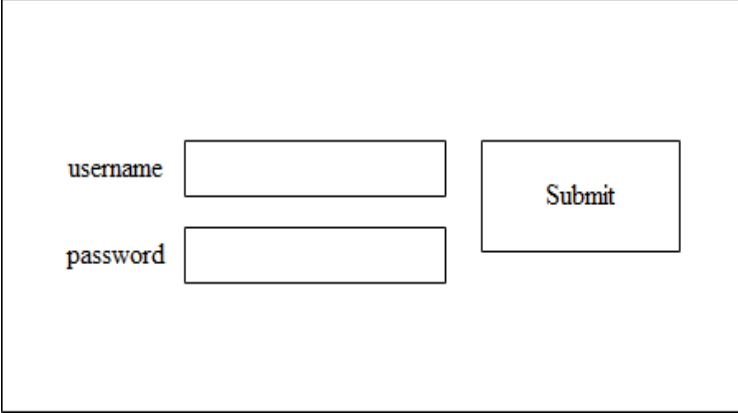


Gambar 5.8. Sequence Diagram deteksi

III.4.6. Desain User Interface

Perancangan User Interface merupakan masukan yang penulis rancang guna lebih memudahkan dalam entry data. Entry data yang dirancang akan lebih mudah dan cepat dan meminimalisir kesalahan penulisan dan memudahkan perubahan. Perancangan User Interface tampilan yang dirancang adalah sebagai berikut :

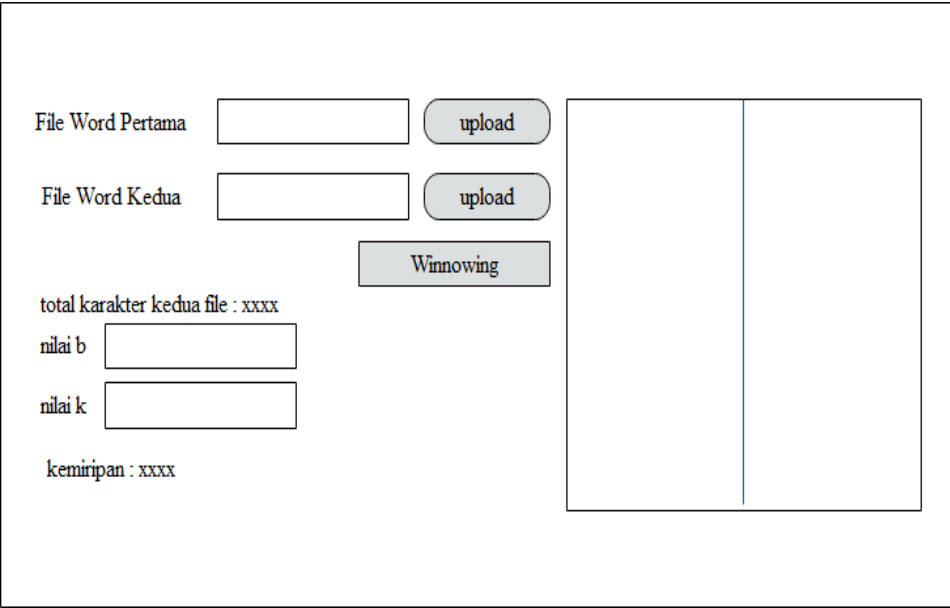
1. Rancangan Form menu login



A login form design within a rectangular border. It features two input fields: the top one is labeled 'username' and the bottom one is labeled 'password'. To the right of these fields is a rectangular button labeled 'Submit'.

Gambar 7.6. Rancangan menu login

2. Rancangan Form menu input dokumen



A document input and detection form design within a rectangular border. It includes two rows of input fields: 'File Word Pertama' and 'File Word Kedua', each followed by an 'upload' button. Below these is a 'Winnowing' button. Further down are three more input fields: 'total karakter kedua file : xxxx', 'nilai b', and 'nilai k'. At the bottom left, there is a label 'kemiripan : xxxx'. On the right side of the form is a large, empty rectangular area, possibly for a preview or output.

Gambar 7.7. Rancangan menu deteksi