Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk menggubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

Team project ©2017
Dony Pratidana S. Hum | Bima Agus Setyawan S. IIP

BAB III

METODE DAN PERANCANGAN SISTEM

3.1 Metodologi Penelitian

Sistem yang dikembangkan berfungsi untuk mengolah berita menjadi suatu ringkasan. Penelitian ini dilakukan dengan menggunakan metode Maximum Marginal Relevance. Sistem mengolah dan memproses suatu berita melalui *input* pada *website*, kemudian menampilkan ringkasan dari proses tersebut. Berikut adalah tahapan-tahapan yang dilakukan dalam penelitian ini.

1. Studi Literatur

   Melakukan pengumpulan informasi berupa teori-teori yang berhubungan dengan penelitian yang dilakukan, seperti teori mengenai *automatic text summarization*, algoritma Maximum Marginal Relevance, dan metode Term Frequency-Inverse Document Frequency (TF-IDF). Selain itu juga dilakukan pencarian data pendukung lainnya berupa referensi dari jurnal, internet, buku, skripsi, dan sumber lainnya. Tahapan ini dilakukan secara berkelanjutan selama proses rancang bangun sistem berlangsung.

2. Perancangan Sistem

   Melakukan perancangan awal terhadap sistem yang dibangun dengan membuat Data Flow Diagram (DFD) dan *flowchart diagram*. Di awali dengan Data Flow Diagram dari keseluruhan sistem yang menerima *input* berupa dokumen teks untuk kemudian diproses untuk menghasilkan *output* berupa ringkasan dari teks tersebut. Dilanjutkan dengan *flowchart diagram* berisi cara kerja sistem secara keseluruhan mulai dari penerimaan *input* dokumen teks sampai dengan *output* berupa hasil ringkasan yang telah diproses oleh sistem.

3. Implementasi

   Melakukan pembangunan aplikasi berbasis *web* dengan menggunakan algoritma Maximum Marginal Relevance. Kemudian dilanjutkan dengan pengujian sistem untuk mengetahui kemampuan kinerja sistem.

4. Uji Coba dan Evaluasi

   Melakukan uji coba terhadap sistem yang telah dikembangkan dan melakukan evaluasi terhadap sistem berdasarkan hasil uji coba yang telah dilakukan.

3.2 Perancangan Sistem

Pembangunan sistem dimulai dengan membuat perancangan awal untuk mengetahui alur kerja sistem. Rancangan sistem yang dibuat adalah sebagai berikut.

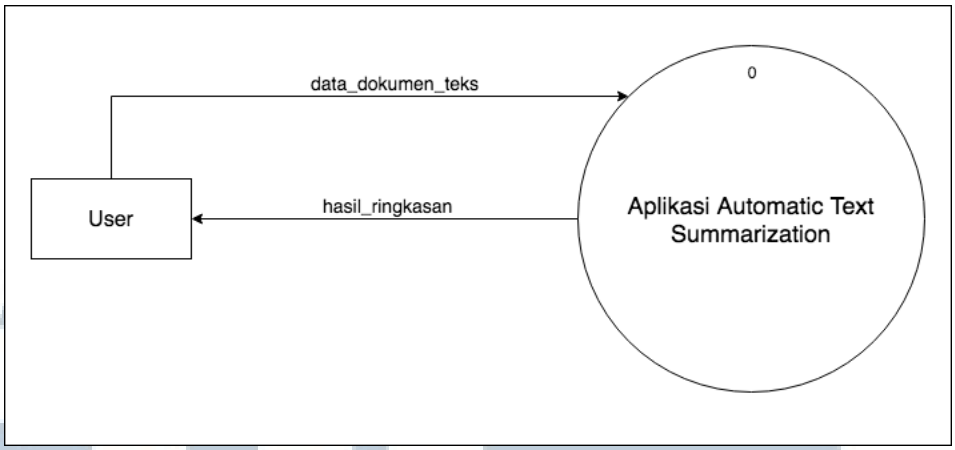
3.2.1 Data Flow Diagram

Data Flow Diagram berfungsi untuk menggambarkan entitas dan data yang dibutuhkan oleh sistem, untuk kemudian diproses mulai dari masuknya data yang dibutuhkan berupa dokumen teks, hingga menghasilkan *output* berupa ringkasan dari teks tersebut. Entitas dan data yang dibutuhkan oleh sistem dijelaskan sebagai berikut.

A. Context Diagram

Sistem menerima data berupa dokumen teks dari *user*, kemudian setelah proses peringkasan teks selesai, *user* menerima *output* berupa hasil ringkasan. Gambar 3.1 merupakan gambar dari *context diagram* atau diagram level 0 dari aplikasi *automatic text summarization*.

Rancang Bangun Aplikasi..., Adrian Renaldi Tanujaya, FTI UMN, 2018



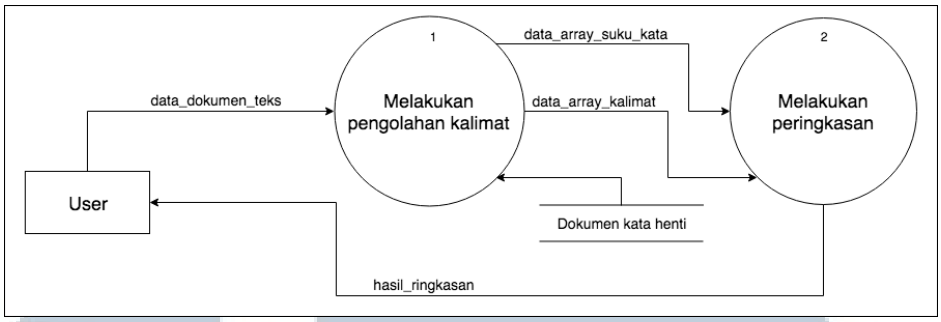
Gambar 3.1 Context Diagram atau Diagram Level 0

B. Diagram Level 1

Pada diagram level 1 ini, sistem melakukan dua buah proses besar berupa pengolahan kalimat dan melakukan proses peringkasan. Pengolahan kalimat merupakan tahap untuk mengubah kalimat-kalimat yang ada pada dokumen teks menjadi sebuah *array* yang berisi kata dasar. Proses pengolahan tersebut juga melibatkan sebuah dokumen berisi kata-kata henti yang digunakan sebagai pembanding dalam salah satu proses pada tahap ini.

Kemudian, *array* berisi kata dasar tersebut diproses ke tahap selanjutnya yaitu tahap peringkasan, setiap kata dasar yang ada dalam *array* dihitung dengan menggunakan metode Maximum Marginal Relevance untuk menghasilkan ringkasan. Gambar 3.2 merupakan diagram level 1 untuk sistem *automatic text summarization*.

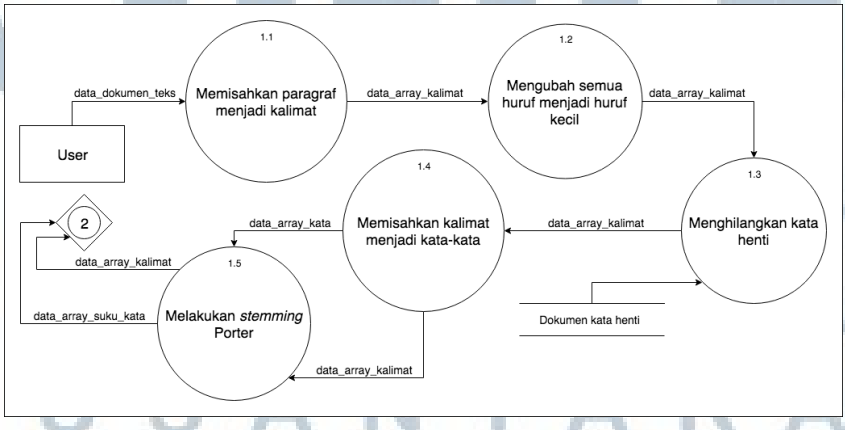
UNIVERSITAS
MULTIMEDIA
NUSANTARA



Gambar 3.2 Diagram Level 1

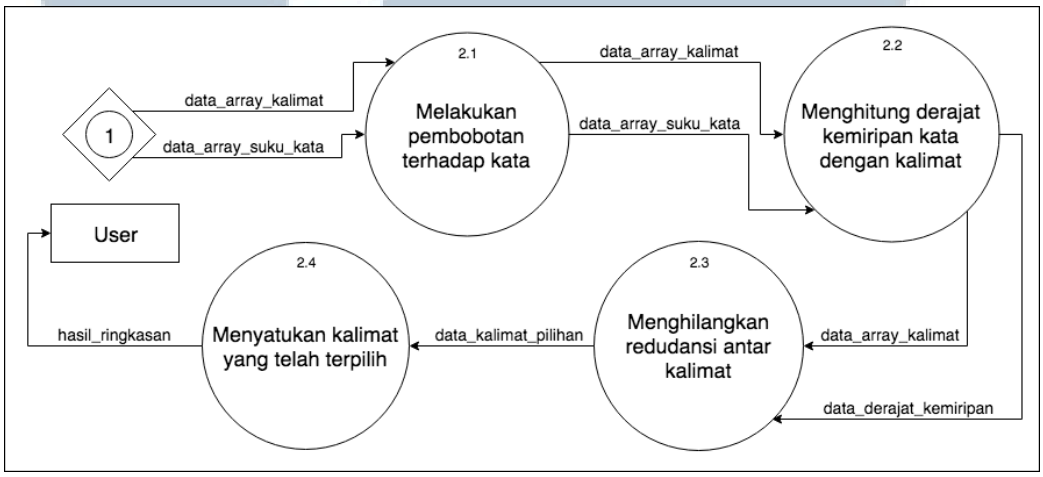
C. Diagram Level 2

Pada diagram level 2, proses pengolahan kalimat dibagi menjadi beberapa proses. Proses pertama adalah pemisahan paragraf menjadi kalimat untuk kemudian dimasukkan ke dalam *array* kalimat. Setiap huruf pada kalimat dalam *array* tersebut, diubah menjadi huruf kecil dan kemudian setiap kata henti yang ditemukan berdasarkan dokumen kata henti, dihilangkan. Proses selanjutnya adalah melakukan pemisahan kalimat menjadi kata-kata dan dimasukkan ke dalam *array* kata untuk menjalani proses *stemming* dengan metode *stemming* Porter. Hasil *stemming* tersebut dikirimkan ke proses selanjutnya untuk dilakukan perhitungan. Gambar 3.3 merupakan gambar dari diagram level 2 untuk proses pengolahan kalimat.



Gambar 3.3 Diagram Level 2 Proses Pengolahan Kalimat

Setelah mendapatkan *array* kata yang merupakan hasil dari proses sebelumnya, masing-masing kata dalam *array* tersebut dihitung bobot masing-masing. Bobot tersebut didapatkan dengan menghitung jumlah kemunculan kata tersebut dalam kalimat. Kemudian, masing-masing kalimat tersebut dihitung derajat kemiripannya. Semakin tinggi derajat kemiripannya, maka kalimat tersebut memiliki makna yang sama. Sebelum kalimat tersebut disatukan kembali menjadi sebuah ringkasan, kalimat yang memiliki makna sama dihilangkan. Gambar 3.4 merupakan gambar dari proses melakukan peringkasan.



Gambar 3.4 Diagram Level 2 Proses Melakukan Peringkasan

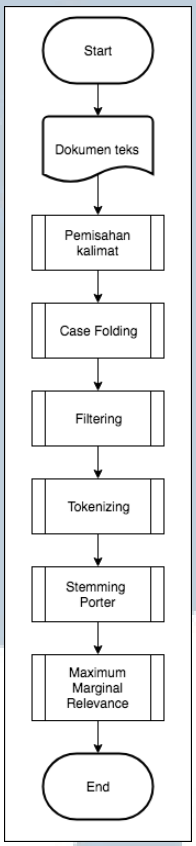
3.2.2 Flowchart

Flowchart berfungsi untuk menggambarkan alur kerja sistem secara keseluruhan. Alur kerja sistem dibagi menjadi beberapa bagian dengan penjelasan sebagai berikut.

A. Flowchart Sistem Peringkas Teks

Sistem menerima input berupa dokumen teks. Setelah itu, sistem membaca isi dokumen tersebut, dan diproses ke dalam 6 tahapan yaitu pemisahan kalimat, *case*

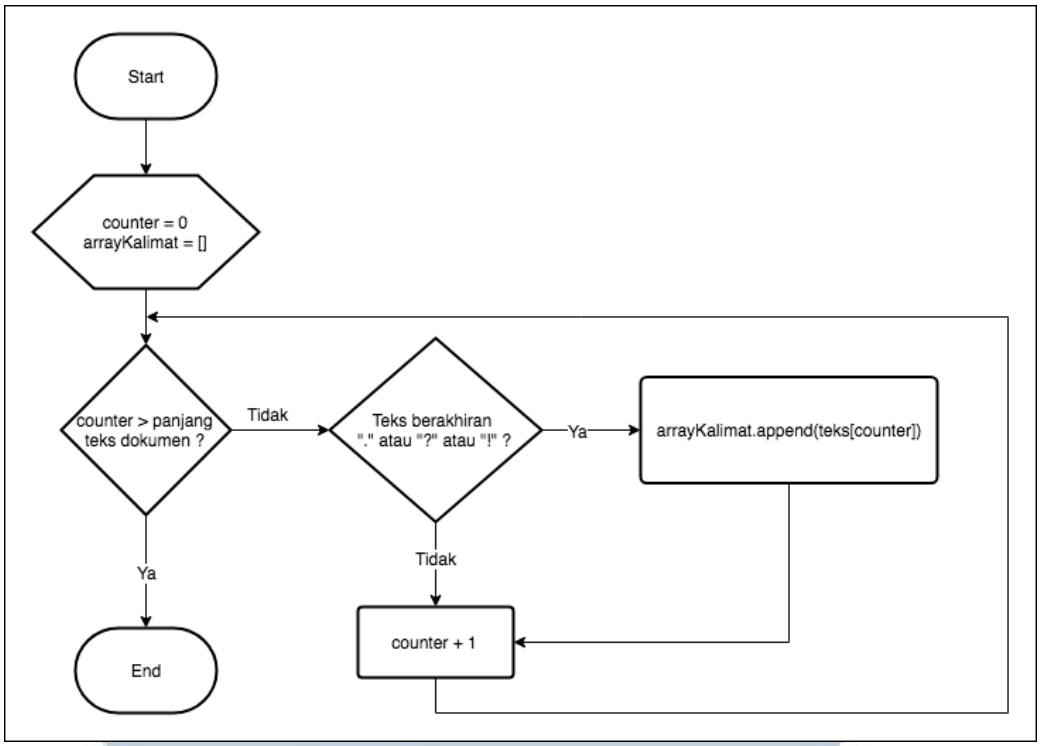
folding, *filtering*, *tokenizing*, *stemming* Porter, dan pengukuran derajat keberagaman dengan Maximum Marginal Relevanve. Hasil dari proses tersebut digunakan untuk mendapat *output* dari sistem. Gambar 3.5 merupakan Flowchart dari sistem.



Gambar 3.5 Flowchart Sistem Peringkas Teks

B. Flowchart Pemisahan Kalimat

Pada proses pemisahan kalimat, teks yang telah dibaca oleh sistem dipisah-pisah menjadi beberapa kalimat dengan cara melakukan iterasi dengan *delimiter* ".", "?", dan "!". Kemudian, hasil dari pemisahan tersebut dimasukkan ke dalam sebuah *array*. Gambar 3.6 merupakan Flowchart dari proses pemisahan kalimat.

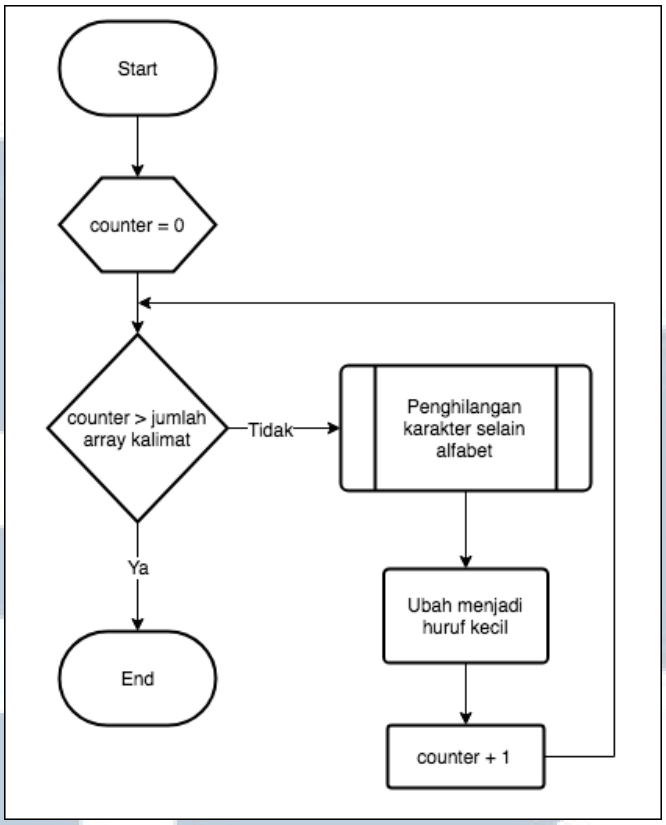


Gambar 3.6 Flowchart Pemisahan Kalimat

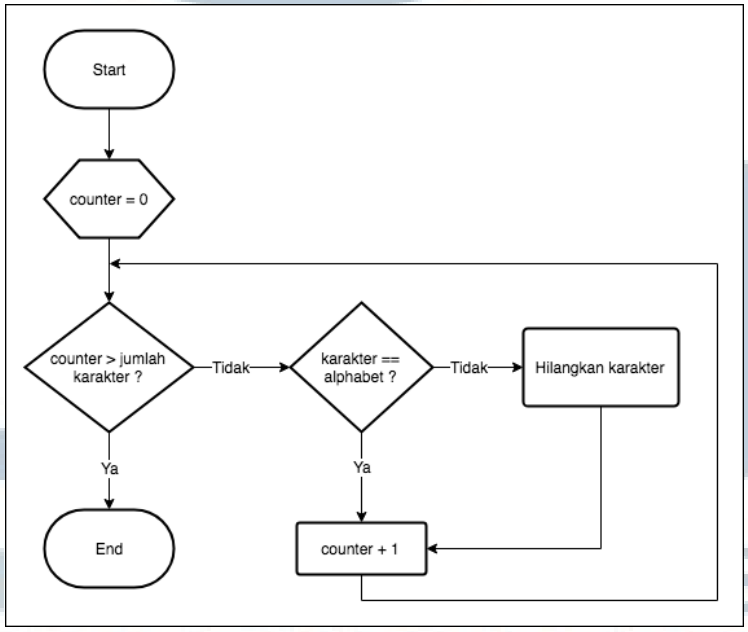
C. Flowchart Case Folding

Pada proses *case folding*, teks yang telah dipisah-pisah ke dalam beberapa kalimat, semua karakter dalam kalimat tersebut diubah menjadi huruf kecil dan dihilangkan karakter yang bukan termasuk ke dalam alfabet. Proses ini mengiterasi semua kalimat yang berasal dari proses sebelumnya. Gambar 3.7 merupakan Flowchart dari proses *case folding* dan Gambar 3.8 merupakan gambar dari proses penghilangan karakter selain alfabet yang ada di dalam proses *case folding*.

UNIVERSITAS
MULTIMEDIA
NUSANTARA



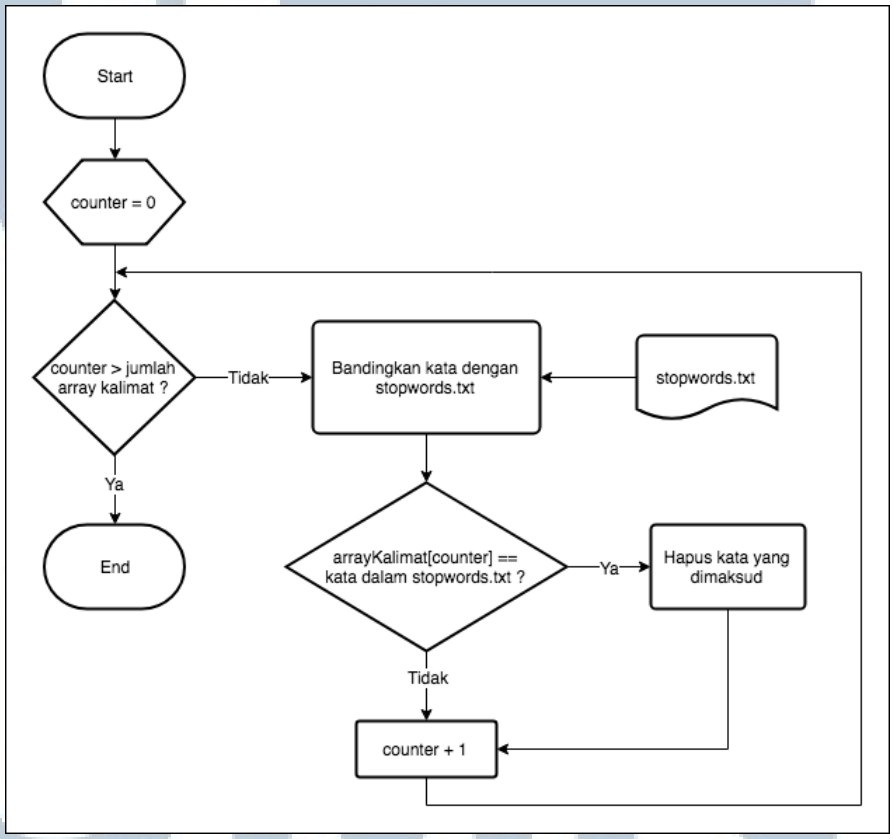
Gambar 3.7 Flowchart Case Folding



Gambar 3.8 Flowchart Penghilangan Karakter Selain Alfabet

D. Flowchart Filtering

Pada proses *filtering*, semua kata henti yang ada dalam kalimat-kalimat dari proses sebelumnya dihilangkan. Kata henti yang menjadi pembanding untuk kata-kata dalam kalimat berasal dari dokumen teks lain yaitu stopwords.txt. Contoh dari kata henti yang dimaksud adalah "jika", "di", "ke", dan lain sebagainya. Jika pada kalimat yang diproses tidak ditemukan kata henti, iterasi berlanjut untuk kalimat berikutnya. Gambar 3.9 merupakan Flowchart untuk proses *filtering*.

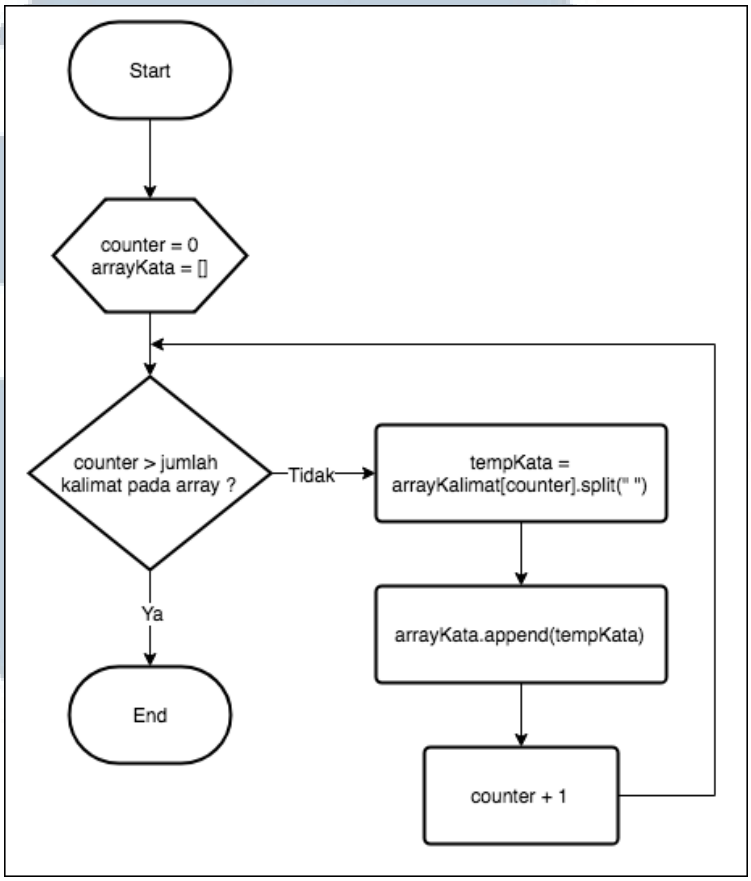


Gambar 3.9 Flowchart Filtering

E. Flowchart Tokenizing

Pada proses *tokenizing*, setiap kalimat yang berasal dari proses sebelumnya dipisahkan menjadi kata-kata berdasarkan spasi antar kata. Kemudian, kata-kata

yang dihasilkan dimasukkan ke dalam sebuah array untuk digunakan pada proses selanjutnya. Gambar 3.10 merupakan Flowchart untuk proses *tokenizing*.

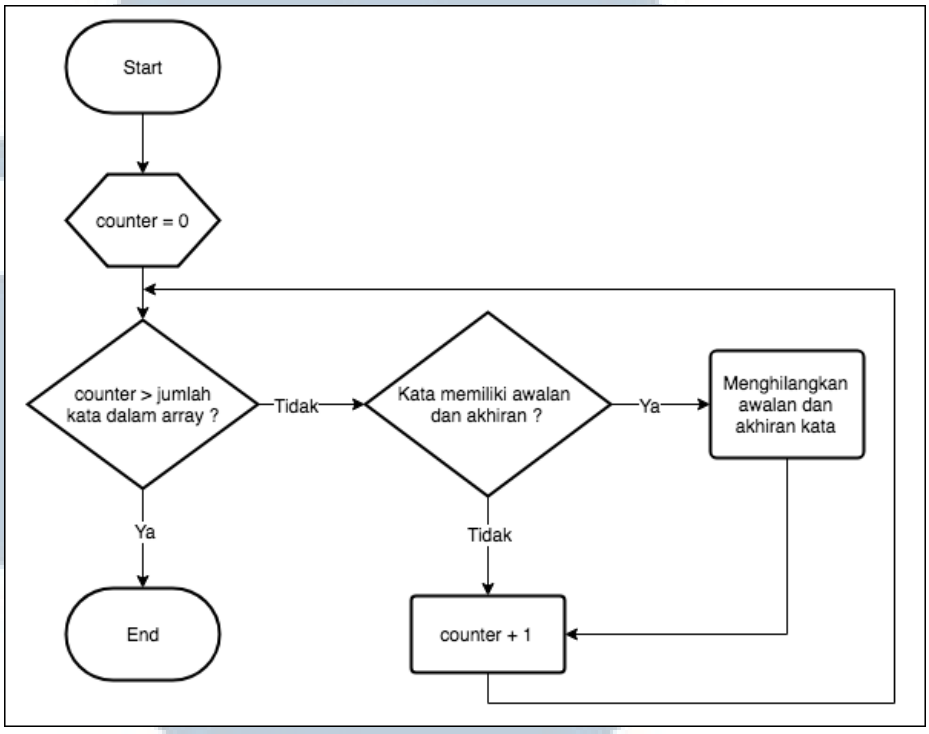


Gambar 3.10 Flowchart Tokenizing

F. Proses Stemming Porter

Pada proses *stemming* Porter, setiap kata yang dihasilkan dari proses sebelumnya, diuraikan berdasarkan kata dasarnya. Proses ini menghapus awalan dan akhiran kata berdasarkan kondisi tertentu. Pada setiap iterasi untuk memproses masing-masing kata, ada beberapa tahap proses eliminasi awalan dan akhiran kata dengan kondisi yang berbeda-beda. Kata dasar yang dihasilkan bukan merupakan kata dasar yang berdasarkan kaidah bahasa, namun kata dasar yang dihasilkan adalah kata dasar yang dapat mempercepat proses pencarian informasi, dalam hal

ini untuk menghasilkan ringkasan teks. Gambar 3.11 merupakan Flowchart proses *stemming* Porter.

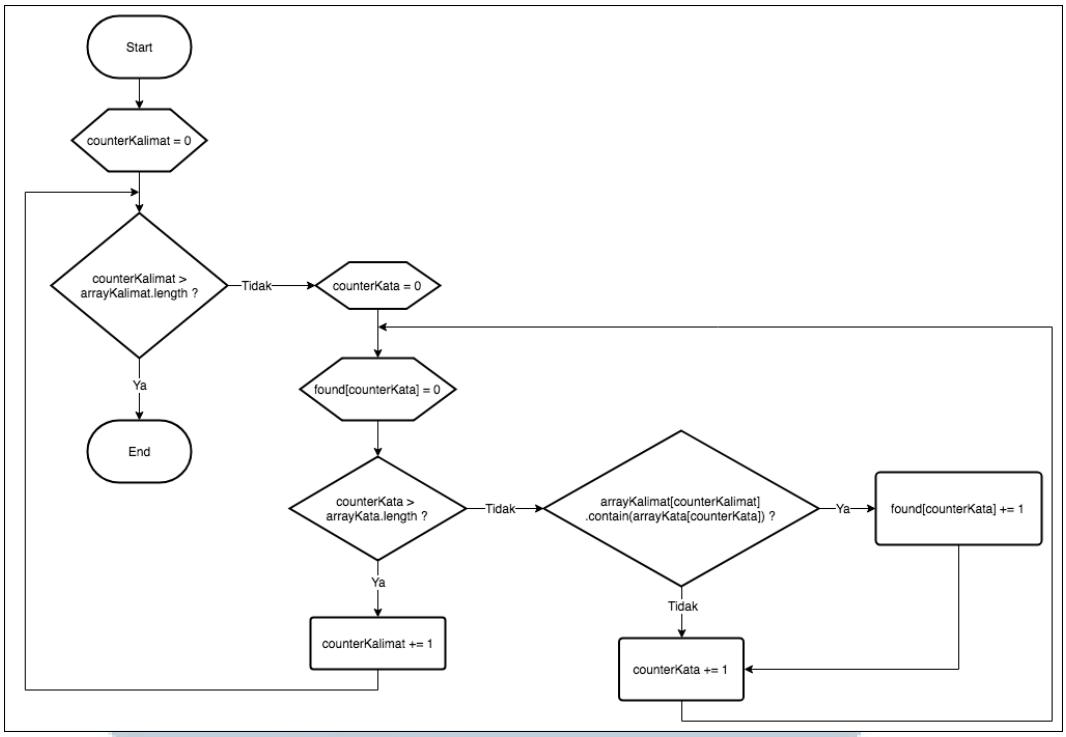


Gambar 3.11 Flowchart Stemming Porter

G. Flowchart Maximum Marginal Relevance

Proses perhitugan Maximum Marginal Relevance dimulai dengan pembuatan vektor antara semua kalimat yang ada dalam dokumen dan *array* kata yang merupakan hasil dari proses sebelumnya. Kemudian, untuk menghitung banyaknya suatu kata di dalam sebuah kalimat, dilakukan proses penghitungan Term Frequency. Gambar 3.12 merupakan Flowchart dari proses penghitungan Term Frequency.

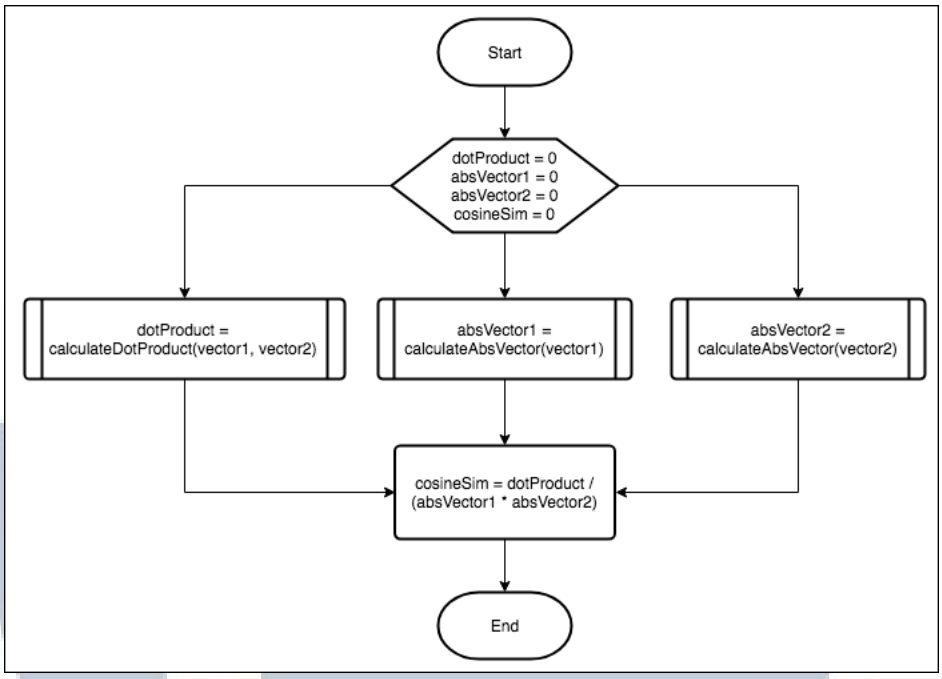
UNIVERSITAS MULTIMEDIA NUSANTARA



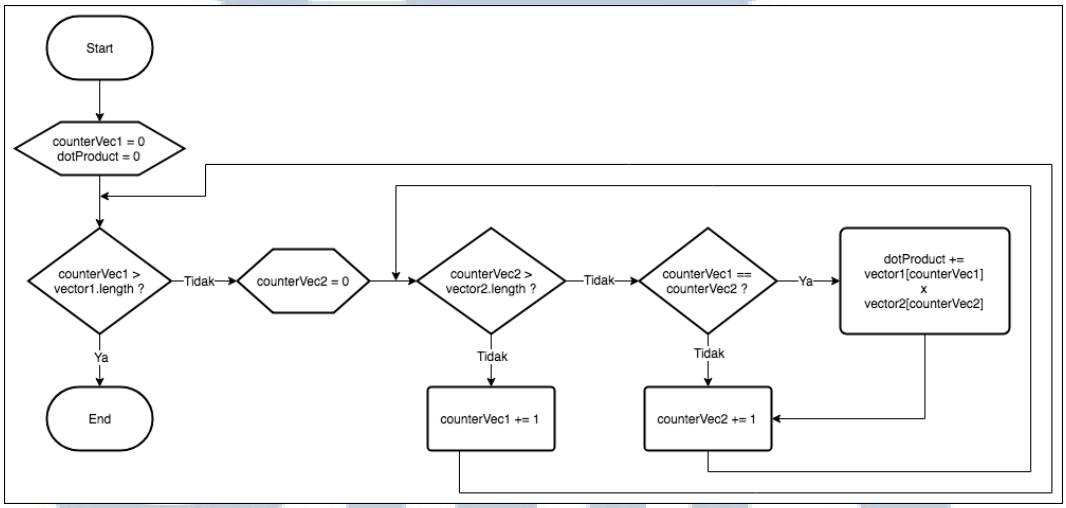
Gambar 3.12 Flowchart Penghitungan Term Frequency

Setelah mendapatkan Term Frequency dari masing-masing kata dalam kalimat, langkah selanjutnya adalah mencari derajat kemiripan antar kalimat dengan menggunakan persamaan *cosine similarity*. Pencarian derajat kemiripan tersebut adalah dengan membagi nilai *dot product* dari *vector* yang berisi Term Frequency antar kalimat dengan nilai perkalian akar pangkat dua dari jumlah Term Frequency kuadrat. Gambar 3.13 merupakan Flowchart dari pencarian derajat kemiripan. Kemudian, Gambar 3.14 dan 3.15 merupakan Flowchart dari pencarian nilai *dot product* antar kalimat dan Flowchart dari pencarian akar pangkat dua dari jumlah Term Frequency kuadrat.

UNIVERSITAS MULTIMEDIA NUSANTARA

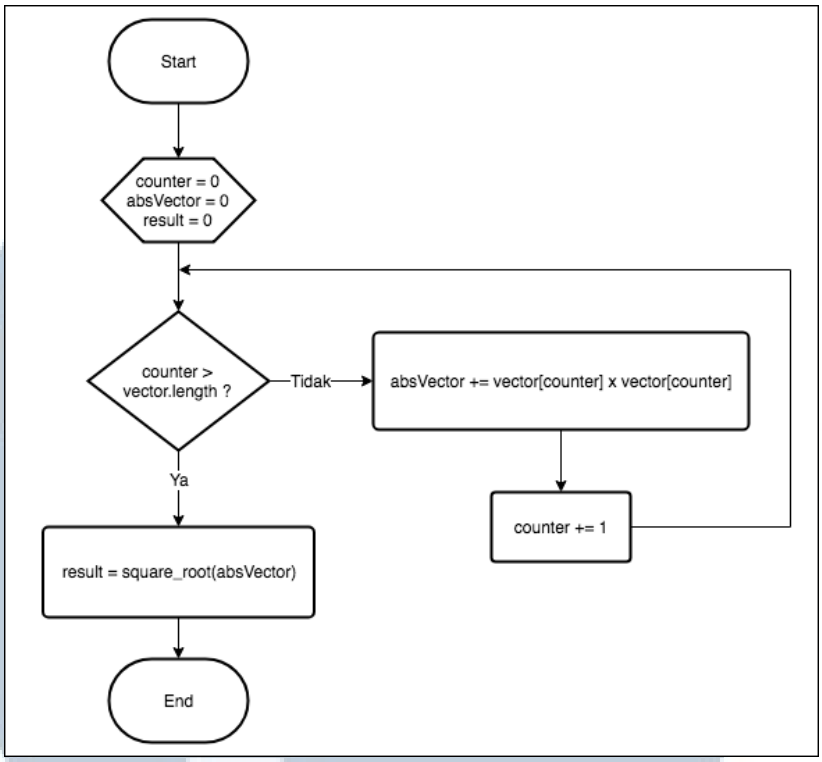


Gambar 3.13 Flowchart Pencarian Derajat Kemiripan



Gambar 3.14 Flowchart Pencarian Dot Product Antar Kalimat

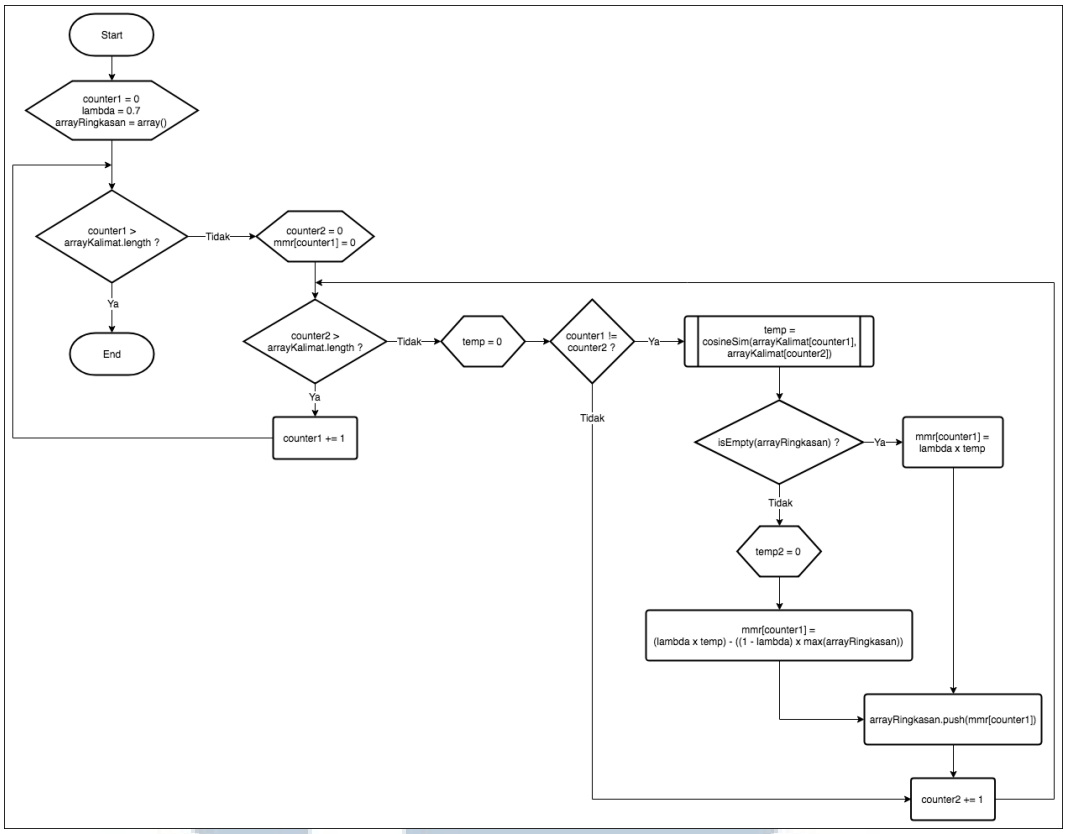
UNIVERSITAS
MULTIMEDIA
NUSANTARA



Gambar 3.15 Flowchart Pencarian Akar Pangkat Dua dari Jumlah Term Frequency Kuadrat

Kemudian, setelah mendapatkan semua derajat kemiripan masing-masing kalimat, proses dilanjutkan dengan mencari nilai Maximum Marginal Relevance dari masing-masing kalimat tersebut. Gambar 3.16 merupakan Flowchart dari pencarian nilai Maximum Marginal Relevance setiap kalimat.

UNIVERSITAS
MULTIMEDIA
NUSANTARA

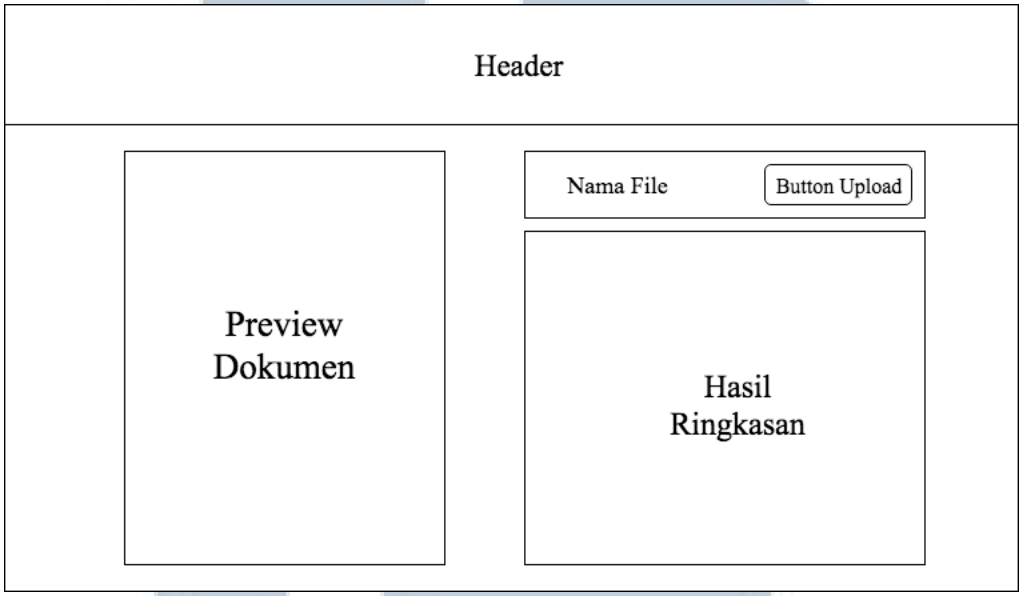


Gambar 3.16 Flowchart Pencarian Nilai Maximum Marginal Relevance

3.2.3 Rancangan Antar Muka

Rancangan antar muka dari aplikasi *automatic text summarization* terdiri dari 1 halaman yang dibagi menjadi 3 jenis proses. Rancangan antar muka pertama merupakan antar muka untuk aplikasi *automatic text summarization* pertama kali dijalankan pada *browser*. Antar muka yang disediakan terdiri dari 4 bagian. Bagian pertama adalah *header*, yang merupakan tempat untuk menunjukkan judul dari aplikasi *automatic text summarization*. Bagian kedua adalah bagian *preview* dokumen, yang merupakan tempat untuk menunjukkan hasil dari unggahan dokumen yang dilakukan oleh *user*. Kemudian pada bagian ketiga, yaitu bagian untuk *user* melakukan unggahan dokumen. Pada bagian tersebut terdapat *field* untuk menunjukkan nama dokumen yang telah diunggah dan tombol untuk

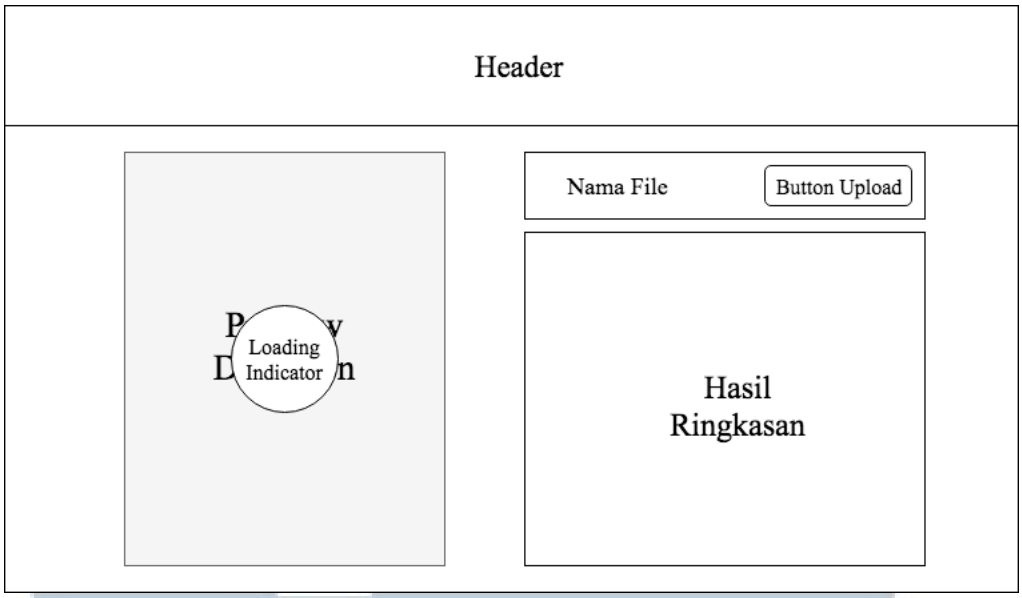
melakukan unggahan dokumen. Bagian terakhir adalah tempat untuk menampilkan hasil ringkasan yang telah melalui proses. Gambar 3.17 merupakan gambar dari rancangan antar muka aplikasi *automatic text summarization*.



Gambar 3.17 Rancangan Antar Muka Aplikasi Automatic Text Summarization

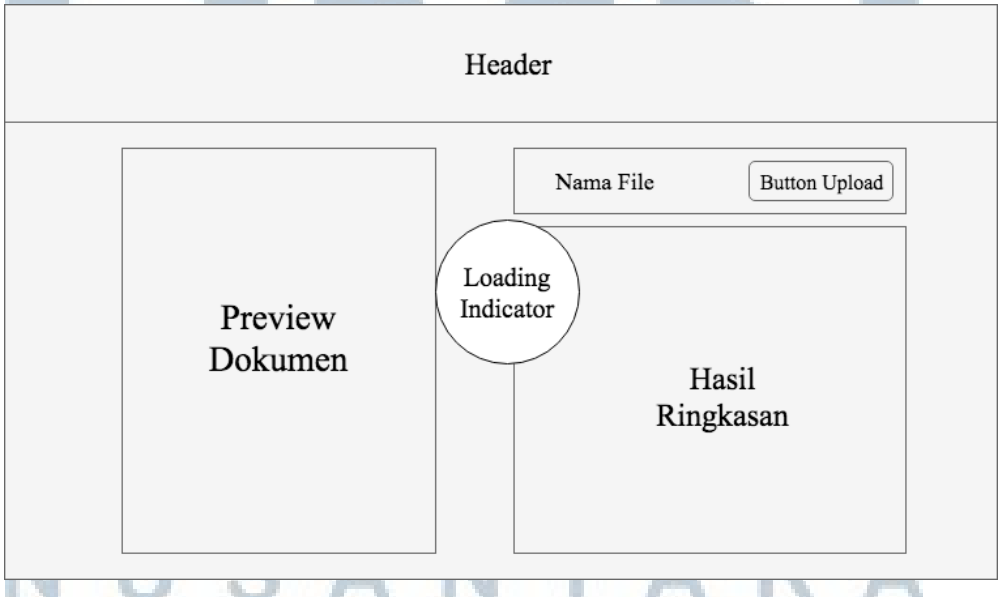
Rancangan antar muka kedua menunjukkan antar muka ketika proses *upload* sedang berjalan dan menunggu dokumen untuk ditampilkan pada bagian *preview* dokumen. Selama proses tersebut berlangsung, ada sebuah *loading indicator* yang menutupi bagian *preview* dokumen hingga proses *upload* dan penampilan dokumen selesai. Gambar 3.18 merupakan rancangan antar muka untuk proses *upload* dan penampilan dokumen.

UNIVERSITAS
MULTIMEDIA
NUSANTARA



Gambar 3.18 Rancangan Antar Muka Proses Upload dan Penampilan Dokumen

Rancangan antar muka ketiga menunjukkan antar muka ketika proses peringkasan dokumen berlangsung. Selama proses berlangsung, muncul sebuah *loading indicator* berputar yang menutupi halaman untuk menunjukkan bahwa proses peringkasan sedang berlangsung. Gambar 3.19 merupakan rancangan antar muka ketika proses peringkasan sedang berlangsung.



Gambar 3.19 Rancangan Antar Muka ketika Proses Peringkasan Berlangsung