# NASKAH PUBLIKASI

PROYEK TUGAS AKHIR

**IMPLEMENTASI CLASTERING NILAI MAHASISWA**

**MENGGUNAKAN METODE MAX**-**MAX ROUGHNESS**

Program Studi Teknik Informatika

Fakultas Teknologi Informasi dan Elektro

Diajukan oleh

**OKTORA KEVIN ARIGI**

**5130411378**

**PROGRAM STUDI TEKNIK INFORMATIKA**

**FAKULTAS TEKNOLOGI INFORMASI DAN ELEKTRO**

**UNIVERSITAS TEKNOLOGI YOGYAKARTA**

**2017**

Naskah Publikasi

**IMPLEMENTASI CLASTERING NILAI MAHASISWA**

**MENGGUNAKAN METODE MAX**-**MAX ROUGHNESS**

Disusun Oleh :

**Oktora Kevin Arigi**

**5130411378**

Telah disetujui oleh pembimbing

Pembimbing

**Suhirman, Ph.D.** Tanggal :

Implementasi Clastering Nilai Mahasiswa

Menggunakan Metode Max-Max Roughness

Oktora Kevin Arigi

*Program Studi Teknik Informatika,Fakultas Teknologi Informasi dan Elektro
Universitas Teknologi Yogykarta*

*Jl. Ringroad Utara Jombor Sleman Yogyakarta*

*E-mail :* *oktora.kevin.arigi@gmail.com*

# **ABSTRAK**

*Klasterisasi teks mempunyai salah satu permasalahan utama dalam mengklasifikasikan teks yang mempunyai sifat uncertain atau sulit dikategorikan. Proses dalam klasterisasi data, banyak metode dan algoritma yang bisa digunakan, seperti algoritma K-means, C-means dan masih banyak yang lainnya. Namun metode tersebut hanya metode dengan data numerik. Tujuan klasterisasi adalah mengelompokkan data ke dalam suatu klaster, sehingga data pada suatu klaster memiliki tingkat kemiripan yang maksimum dan data antar kluster memiliki kemiripan yang minimum. Pengelompokkan data kedalam beberapa klaster dapat digunakan untuk mengambil keputussan. Hasil keputusan salah satunya bisa digunakan sebagai acuan dalam proses menentukan kebutuhan pembelajaran sehingga proses belajar lebih efektif. Salah satu metode yang dapat digunakan dalam proses clustering dengan data nilai mahasiswa adalah Max-Max Roughness(MMR). MMR menggunakan teori rough set, dari data yang disajikan menggunakan teori rough set dipetakan kedalam tabel keputusan dengan memperkenalkan atribut keputusan. Value suatu atribut, yang memiliki nilai mean roughness maksimum, akan digunakan sebagai dasar dalam membentuk klaster. Dalam penerapan metode MMR data yang digunakan diambil dari mahasiswa Teknik Informatika Universitas Teknologi Yogyakarta. Berdasrkan hasil penerapan metode MMR dengan data nilai mahasiswa Teknik Informatika didapatkan hasil bahwa metode MMR dapat digunakan dalam pengelompokkan data bersifat kategori. Hasil menunjukkan bahwa atribut yang dominan pada atribut lain dapat ditentukan.*

**Kata kunci** : Nilai, MMR, *Clustering*, Teori *Rough Set*

### 1. PENDAHULUAN

*Clustering* adalah metode penganalisaan data, yang sering dimasukkan sebagai salah satu metode data *mining*. Tujuannya adalah untuk mengelompokkan data dengan karakteristik ke suatu ‘wilayah’ yang sama dan data dengan karakteristik yang berbeda ke 'wilayah' yang lain.

Permasalahan yang ditemui adalah menentukan suatu atribut dari data nilai mahasiswa yang akan digunakan sebagai acuan dalam pengelompokkan data. Selain itu permasalahan lain dalam mengklasifikasikan nilai yang berupa teks adalah mempunyai sifat *uncertain* atau sulit dikategorikan. Menggunakan metode yang ada dalam proses *clustering*, maka pengelompokkan data yang terbentuk dalam suatu klaster memiliki tingkat kemiripan yang maksimum dan data antar kluster memiliki kemiripan yang minimum. Tujuan *clustering* untuk menghasilkan pengelompokkan objek yang mirip satu sama lain dalam kelompok. Selain itu dapat menangani *uncertainty* pada saat proses *clustering* data. Metode yang dapat digunakan dalam proses *clustering* data sering diterapkan pada data yang bersifat numerik. Beberapa metode dengan data numerik adalah K-means, C-means dan masih banyak yang lainnya.

Metode dalam proses *clustering* dengan data yang bersifat ketegori, ada beberapa yaitu *Total Roughness* (TR), *Max-Max Roughness* (MMR) dan lain-lain. Metode MMR merupakan metode perbaikan dari metode TR, maka dalam proses *clustering* nilai mahasiswa, metode MMR yang akan digunakan. Metode MMR menerapkan penggunaan *Rough Set Theory* (RST). RST menggunakan konsep *lower approximation* dan *upper approximation* yang akan digunakan untuk mendapatkan nilai *roughness* pada tiap-tiap atribut terhadap atribut lainnya. Selanjutnya metode MMR mencari nilai maksimum *mean roughness* pada tiap-tiap atribut berdasarkan seluruh atribut lainnya, nilai *mean roughness* tertinggi, akan dijadikan atribut acuan dalam proses pengelompokkan data. Batasan masalah dalam penelitian adalah data yang digunakan yaitu data nilai mahasiswa Teknik Informatika Universitas Teknologi Yogyakarta dan atribut yang digunakan dalam proses perhitungan menggunakan metode MMR adalah nilai mata kuliah Teknik Informatika. Sehingga tujuan yang ingin dicapai dari hasil penelitian adalah dapat mengimplementasikan algoritma MMR dalam proses *clustering* nilai mahasiswa dan dapat mengevaluasi nilai yang dihasilkan dari metode MMR dalam proses *clustering* nilai mahasiswa. Permasalahan yang ditemui dalam penelitian ini adalah mendapatkan nilai *mean roughness* dari nilai mahasiswa, menentukan atribut mata kuliah dan menampilkan hasil dalam proses perhitungan menggunakan metode MMR. Selain itu manfaat yang ingin dicapai dari hasil penelitian adalah dapat mempercepat dan mempermudah dalam menentukan atribut yang dominan dalam pengelompokkan nilai mahasiswa, memberikan informasi kepada dosen untuk meningkatkan keefektifan dalam proses mengajar sehingga target nilai dan pemahaman ilmu suatu mata kuliah dapat dicapai, dapat mengetahui tingkat keberhasilan dan keefektifan dosen dalam proses memberikan pengajaran kepada mahasiswa, mengurangi fungsi kuesioner yang diberikan kepada mahasiswa dalam memberikan penilaian terhadap proses mengajar dosen dan sistem dapat diimplementasikan pada akademik kampus dalam menentukan atribut mata kuliah yang dominan.

### 2. LANDASAN TEORI

Metode *clustering* sangat tidak asing dalam bidang data *mining* karena banyak yang sudah menelitinya. Metode MMR merupakan bagian dari data *mining* dalam mendapatkan informasi baru yang terdapat dalam bongkahan basis data. Berikut referensi yang menjadi dasar acuan untuk pembuatan sistem ini.

Referensi [1] penelitian dengan judul Klasterisasi Teks Menggunakan Metode Max-Max Roughness (MMR) Dengan Pengayaan Similaritas Kata. Penelitian tersebut membahas bagaimana mengklasifikasikan jenis teks yang mempunyai sifat uncertainty atau sulit dikategorikan pada data berdimensi yang tinggi dan menyebar. Hasil yang diperoleh dari penggunaan metode MMR dapat menghasilkan kualitas cluster yang lebih baik dibandingkan dengan metode FW-Kmeans.

Referensi [4] penelitian dengan judul Rough Set Approach for Categorical Data Clustering. Penelitian tersebut lebih fokus membahas pada rough set theory untuk kategorikal clustering data. Penelitian ini membahas perbandingan TR (*Total Roughness*) dan MMR (*Minimum-Minimum Roughness*), juga menerapkan metode MADE (Maksimal Atribut *Depedency*). Hasil perbandingan menggunakan metode TR dan MMR didapatkan kesimpulan perhitungan dengan TR dan MMR sama, dari segi waktu didapatkan hasil yang sama juga. Untuk mencapai kompleksitas perhitungan yang lebih rendah dalam memilih atribut partisi menggunakan metode MMR. Metode MRR merupakan perluasan dari metode TR.

Referensi [2] penelitian dengan judul Analisis Implementasi Algoritma Min-Min Roughness (MMR) Berbasis Rough Set Theory Dalam Clustering Pada Data Kategorikal. Penelitian tersebut membahas bagaimana menerapkan algoritma MMR pada data kategorikal. Permasalahan yang dihadapi adalah bagaimana penerapan algoritma MMR ini dalam menyelesaikan permasalahan clustering data kategorikal. Hasil yang diperoleh dari penerapan algoritma MMR, selanjutnya terbentuk cluster-cluster yang akan di analisis tiap *cluster* tersebut dengan menggunakan metode *purity* dan berdasarkan objek-objek yang terdapat dalam *cluster* tersebut. Dapat disimpulkan bahwa algoritma MMR dapat melakukan clustering pada data kategorikal dengan nilai *purity* diatas 0.7

Referensi [11] penelitian judul An Application Of Rough Set Theory To Cluster Student Assessment At Universities. Penelitian tersebut membahas penerapan rough set theory dengan pendengkatan atau metode Max-Max Roughness (MMR). Permasalahan yang diangkat adalah bagaimana menemukan pola dan struktur dasar data untuk membuat proses belajar lebih efektif. Sehingga sistem pendidikan melahirkan konten disesuaikan dengan kebutuhan pengguna tertentu. Personalisasi yang memadai membutuhkan domain pembelajaran yang akan dijelaskan secara eksplisit dalam detail tertentu. Kesimpulan dari hasil penelitian tersebut menunjukkan bahwa atribut dominan pada atribut lain dapat ditentukan, atribut lainnya dapat diabaikan, sehingga proses penilaian dan penyediaan rekomendasi bisa dilakukan lebih cepat.

Referensi [12] penelitian dengan judul Min-Min Roughness (MMR) untuk Klasterisasi Data Kategori Studi Kasus Kecemasan Belajar pada Mahasiswa Tahun Pertama dan Keempat di UAD. Penelitian tersebut membahas penerapan metode MMR dalam mengklasifikasikan sumber kecemasan belajar pada mahasiswa. Kesimpulan dari hasil penelitian tersebut adalah bahwa metode MMR dapat digunakan dalam mengklasterisasi semua sumber kecemasan belajar pada mahasiswa.

Dari beberapa penelitian yang telah ada, dapat disimpulkan bahwa dalam penerapan metode MMR dalam proses *clustering* dapat menangani *uncertainty* pada saat proses *clustering* data. Sehingga dari hasil penelitian yang telah dilakukan bagaimana mengimplementasikan *clastering* nilai mahasiswa menggunakan metode MMR sehingga dapat mempercepat dan mempermudah dalam *clustering* nilai mahasiswa.

**2.1** **ROUGHT SET THEORY**

*Rought set theory*, diperkenalkan oleh Zdzislaw Pawlak pada awal 1980-an, merupakan teori pendekatan untuk ketidak jelasan. *Rought set theory* adalah suatu metode atau alat bantu dalam analisis data yang disajikan dalam bentuk table [2]. *Rough Set* merupakan salah satu teknik data *mining* yang digunakan untuk menangani masalah *Uncertainty*, *Imprecision* dan *Vagueness* dalam aplikasi *Artificial Intelligence* (AI). *Rough set* merupakan teknik yang efisien untuk *Knowledge Discovery in Database* (KDD) dalam tahapan proses dan Data Mining [6]. Permasalahan yang *uncertainty* disini adalah ketidak pastian pada hasil pengambilan keputusan yang disebabkan adanya *imprecise knowledge*, yaitu data atau objek yang nilai kebenarannya dianggap kurang tepat. *Rough set* akan digunakan untuk mendapatkan nilai roughnesss.

Berikut ini adalah *nomenclature* yang digunakan :

$U$ : *Universal* atau himpunan dari seluruh objek ($x\_{1},x\_{2}$,...)

$X$ : Subset dari himpunan seluruh objek $\left(X⊂U\right)$

$x\_{i}$ : Objek yang dimiliki oleh subset himpunan seluruh objek $x\_{i}\in X$

$A$ : Himpunan seluruh atribut

$a\_{i}$ : Atribut yang dimiliki oleh himpunan seluruh atribut

V($a\_{i}$) : Himpunan *value* dari atribut $a\_{i}$

$B$ : Subset tidak kosong dari himpunan seluruh atribut $\left(B⊆A\right)$

$\overline{X\_{B}}$ : *Lower Approximation* dari $X$ terhadap $B$

$\overline{X\_{B}}$ : *Upper Approximation* dari $X$ terhadap $B$

$R\_{a\_{i}}\left(X\right)$ : *Roughness* terhadap $\left\{a\_{i}\right\}$

Ind$\left(B\right)$ : *Indiscernibility Relation*

$\left[X\_{i}\right]\_{Ind\left(B\right)}$ : *Equivalence class* $x\_{i}$ pada relasi Ind$\left(B\right)$, juga disebut *elementary* *set* di $B$

***Tabel 1*** *Contoh Data Mobil*

|  |  |  |  |
| --- | --- | --- | --- |
| **Nama Mobil** | **Warna** | **Jumlah Ban** | **Tipe** |
| A1 | Hitam | 4 | Minibus |
| A2 | Putih | 4 | Pick Up |
| A3 | Putih | 6 | Truk |
| A4 | Merah | 6 | Truk |
| A5 | Biru | 4 | Minibus |
| A6 | Hitam | 6 | Truk |
| A7 | Merah | 10 | Truk |

#### a. Indiscernibility Relation

*Indiscernibility relation* adalah suatu kondisi dimana suatu objek tidak dapat dibedakan (*indiscernible*) terhadap objek yang lain pada suatu atribut ataupun lebih [2]. Ind$\left(B\right)$ adalah relasi yang terdapat pada $U$ dimana diberikan dua objek $x\_{i}$ dan $x\_{j}$ dimana $x\_{i},x\_{j}\in U$ dapat dibedakan oleh himpuanan atribut $B$ dalam $A$ jika dan hanya jika $a\left(x\_{i}\right)=a\left(x\_{j}\right)$ untuk setiap $a\in B$. Jadi ($x\_{i},x\_{j}\in Ind\left(B\right)$ jika dan hanya jika $∀ a\in B$ dimana $B⊆A, a\left(x\_{i}\right)=a\left(x\_{j}\right)$ (Parmar, D., dkk, 2007).

A1, A6 $\in $ (Warna=Hitam)

A3, A4, A6, A7 $\in $ (Tipe=Truk)

Objek A1 dan A6 tidak dapat dibedakan jika dilihat berdasarkan atribut warna, karena kedua objek memiliki *value* yang sama untuk atribut tersebut, yaitu hitam. Begitu juga dengan objek A3, A4, A6, A7 apa bila dilihat berdasarkan atribut tipe dimana seluruh objek tersebut memiliki *value* yang sama, yaitu truk.

b. Equivalence Class

Setelah tahap *indiscernibility relation* ditentukan selanjutnya memasukkan *indiscernibility relation* dalam satu himpunan *equivalence class*. *Equivalence class* adalah himpunan berisi objek-objek yang memiliki *indiscernibility relation* pada tiap *value* pada suatu atribut [2]. Diberikan Ind$\left(B\right)$, *equivalence class* $\left(\left[x\_{i}\right]\_{Ind\left(B\right)}\right) $adalah masing-masing himpunan objek dari $x\_{i}$ yang mempunyai *value* yang sama pada himpunan atribut $B$ yang terdiri dari *equivalence class* $\left(\left[x\_{i}\right]\_{Ind\left(B\right)}\right)$. Dikenal juga dengan istilah *elementary set* terhadap $B$ [8]. Data di tabel 2.1, contoh *elementary set* dari atribut warna adalah sebagai berikut :

$X$(Warna=hitam)={A1, A6}; $X$(Warna=putih)={A2, A3};

$X$(Warna=merah)={A4, A7};

$X$(Warna=biru)={A5};

*Elementary set* dari atribut Tipe adalah

{(A1, A5),A2,(A3,A4,A6,A7)}

A1, A5 memiliki *value* yang sama yaitu minibus, A2 *value*-nya adalah *Pick Up* merupakan objek yang tidak memiliki *value* yang sama dengan objek yang lain pada atribut tersebut. Value A3, A4, A6, A7 adalah truk.

c. Lower Approximation

*Lower Approximation* adalah himpunan objek yang pasti termasuk dalam konsep [2]. Diberikan himpunan atribut $B$ di $A$, himpunan dari objek $X$ di $U$, *lower approximation* dari $X$ didefinisakan sebagai *union* dari semua *elementary set* yang terdapat dalam $X$ [8].

$$\overline{X\_{B}}=∪\left\{x\_{i}|\left[x\_{i}\right]\_{Ind\left(B\right)}⊆X\right\}$$

d. Upper Approximation

*Upper approximation* adalah himpunan objek yang mungkin termasuk dalam konsep [2]. Diberikan himpunan atribut $B$ di $A$, himpunan objek $X$ di $U$, *upper approximation* dari $X$ didefinisakan sebagai *union* dari semua *elementary set* yang mempunyai *non-empty* *intersection* terhadap $X$ (Parmar, D., dkk, 2007).

$$\overline{X\_{B}}=∪\left\{x\_{i}|\left[x\_{i}\right]\_{Ind\left(B\right)}∩X\ne ∅\right\}$$

e. Roughness

*Roughness* adalah rasio kardinalitas dari *lower approximation* dan kardinalitas dari *upper approximation* yang didefinisakan sebagai ketepatan dari estimasi, yang merupakan ukuran dari *roughness* [8].

$$R\_{B}\left(X\right)=\frac{\left|\overline{XB}\right|}{\left|\overline{XB}\right|}$$

**2.2 MAX-MAX ROUGHNESS**

Referensi [8] memperkenalkan metode MMR dalam proses *clustering* dengan menggunakan RST. Metode MMR dapat melakukan perhitungan yang kompleks dengan perhitungan yang lebih rendah dalam memilih atribut partisi. Metode MMR dapat digunakan untuk mengklasterisasi data kategori yang memiliki atribut biner maupun multi nilai. Metode ini menggunakan teori *rough set* untuk memilih atribut keputusan dan mengklasterisasi objek pada semua atribut. Metode MMR memanfaatkan nilai *roughness* dari teori *rough set*, selanjutnya adalah mencari nilai *roughness* maksimum (*max-roughness*) dari *mean roughness* dari setiap atribut. Atribut yang memiliki nilai *roughness* maksium akan dipilih sebagai acuan dalam proses *clustering* [2].

a. Mean Roughness

Referensi [2] *Mean Roughness* adalah nilai rata-rata pada atribut $a\_{i}$ terhadap atribut $\left\{a\_{j}\right\}$. Diberikan $a\_{i}\in A, V\left(a\_{i}\right)$ himpunan *value* dari atribut $a\_{i}, X$ adalah subset dari objek yang mempunyai satu *value* spesifik $a$ dari atribut $a\_{i}$ yaitu $X\left(a\_{i}=α\right), \overline{X\_{a\_{j}}\left(a\_{i}=α\right)}$ mengacu pada *upper approximation* dan $\overline{X\_{a\_{j}}\left(a\_{i}=a\right)}$ adalah *lower approximation* terhadap $\left\{a\_{j}\right\}$, $R\_{a\_{j}}\left(X\right)$ yang didefinisikan sebagai *roughness* dari $X$ terhadap $a\_{j}$, maka [8]

$$R\_{aj}\left(X|a\_{i}=α\right)=\frac{\left|\overline{X\_{a\_{j}}\left(a\_{i}=α\right)}\right|}{\left|\overline{X\_{a\_{j}}\left(a\_{i}=α\right)}\right|}$$

dimana $a\_{i},a\_{j}\in A$ dan $a\_{i}\ne a\_{j}$

$\left|V\left(a\_{i}\right)\right|$ adalah jumlah *value* di atribut $a\_{i}$, maka *mean roughness* pada atribut $a\_{i}$ terhadap$\left\{a\_{j}\right\}$ adalah

$Rough\_{a\_{j}}\left(a\_{i}\right)=\frac{R\_{a\_{j}}\left(X|a\_{i}=α\_{1}\right)+ … R\_{a\_{j}}\left(X|a\_{i}=α\_{\left|V\left(a\_{i}\right)\right|}\right)}{\left|V\left(a\_{i}\right)\right|}$

dimana $a\_{i},a\_{j}\in A$ dan $a\_{i}\ne a\_{j}$

b. Max-Roughness

*Max-Roughness* (MR) adalah nilai maksimum *roughness* dari atribut yang diberikan $a\_{i}\in A, V\left(a\_{i}\right)$ mengacu pada maksimum dari *mean roughness* [8]. Apabila diberikan sejumlah $n$ atribut, MR dari atribut $a\_{i}$ adalah

MR($a\_{i}$)=Max ($Rough\_{a\_{1}}$($a\_{i}$),.. $Rough\_{a\_{j}}$($a\_{i}$)...,

dimana $a\_{i},a\_{j}\in A , a\_{i}\ne a\_{j}, 1\leq i, j\leq n$.

c. Max-Max Roughness

*Max-Max Roughness* (MMR) adalah nilai maksimum dari MR semua atribut yang ada [2].

MMR=Max (MR($a\_{1}$)..., MR($a\_{i}$)...,

dimana $a\_{i}\in A , i$ mulai dari 1 sampai kardinalitas $A$

### 3. METODOLOGI PENELITIAN

Dalam penelitian ini, penulis menggunakan metode penelitian dengan tahapan-tahapan yang sudah disiapkan yaitu :

**3.1. PENGUMPULAN DATA**

Pengumpulan data yang dibutuhkan dalam penelitian adalah data nilai dan data yang berkaitan dengan metode MMR. Data nilai dalam penelitian digunakan untuk menerapkan metode MMR dalam proses *clustering* nilai mahasiswa Teknik Informatika. Berdasarkan sumber data yang dibutuhkan yang akan digunakan pada penelitian proyek tugas akhir ini, maka data yang diperoleh melalui beberaapa metode pengumpulan data, yaitu adalah sebagai berikut :

1. Wawancara

Wawancara dilakukan dengan mahasiswa dan wali dosen Teknik Informatika Universitas Teknologi Yogyakarta. Wawancara dilakukan dengan mahasiswa karena sebagian besar mahasiswa memiliki hasil rekap nilai. Wawancara juga dilakukan dengan dosen wali karena sebagian dosen wali memiliki data nilai mahasiswa. Proses wawancara dan pengambilan nilai dilakukan pada tanggal 13 sampai 26 Mei 2017.

1. Studi Literatur

Pencarian data berupa referensi literatur yang berkaitan dengan *clustering*, metode MMR dan *Rough Set Theory*.

1. Dokumentasi

Metode yang digunakan untuk mengumpulkan data dalam penelitian yaitu dengan cara meminta hasil rekap nilai dari mahasiswa dan hasil rekap KHS dari wali dosen Teknik Informatika Universitas Teknologi Yogyakarta. Jumlah data nilai mahasiswa yang akan diambil adalah 300 orang. Data nilai yang digunakan pada tahap pemilihan atribut nilai yaitu :

1. Nama

Atribut “Nama” merupakan atribut yang digunakan sebagai tanda pengenal dari nilai mahasiswa. Selain itu Nama mahasiswa juga digunakan untuk mengetahui nilai yang diperoleh oleh mahasiswa berdasarkan mata kuliah.

1. Nilai Mata Kuliah

Nilai mata kuliah merupakan atribut tetap. Atribut ini yang akan ditentukan dalam proses *clustering* nilai mahasiswa.

***Tabel 2*** *Atribut Mata Kuliah*

|  |  |
| --- | --- |
| **No** | **Atribut** |
| 1 | Sekuriti Komputer |
| 2 | Jaringan Komputer |
| 3 | Logika Informatika |
| 4 | Sistem Basis Data |
| 5 | Pemograman Terstruktur |
| 6 | Metodologi Pengembangan Sistem |

Pada tabel 2 merupakan atribut mata kuliah yang akan digunakan dalam proses *clustering* nilai mahasiswa. Mata kuliah yang dipilih merupakan mata kuliah yang hampir semua mahasiswa sudah menempuhnya, mata kuliah tersebut juga memiliki tingkat keberagaman nilai yang tinggi sehingga semua perhitungan dan algoritma dalam program dapat berjalan, selain itu nilai mata kuliah tersebut mudah dalam proses memperoleh datanya.

### 3.2. Analisis dan Perancangan

Pada tahap analisa pengembangan sistem beberapa tahap yang perlu diperhatikan dalam pembuatan sistem, yaitu tahap basis data, *input*, proses, dan *output*. Tahap basis data untuk penyimpanan menggunakan MySQL dengan jumlah tabel dua. Pada tahap pengimputan data terdiri dua cara yaitu pertama dengan mengimputkan data nilai mahasiswa secara satu persatu melalui *web*, kedua dengan mengimport data nilai mahasiswa dengan *format* excel. Pada tahap proses, data nilai akan dilakukan proses *clustering* dengan metode MMR untuk menentukan atribut yang dominan. Setelah tahap proses *clustering,* akan ditampilkan nilai-nilai *mean roughness* dari setiap atribut dalam bentuk tabel, dari tebel hasil perhitungan *mean roughness* akan dicari nilai maksimum dari semua atribut, sehingga didapat satu atribut dengan nilai *mean roughness* tertinggi, nilai tertinggi akan dijadikan atribut untuk membentuk kluster-kluster dengan tingkat kemiripan maksimum. Rancangan system yang dibangun akan digambarkan menggunakan beberapa permodelan yaitu *flowchart*, *use case*, *sequence diagram*, dan *activity diagram*.

1. Flowchart



***Gambar 1*** *Flowchart Metode MMR*

Gambar 1 menunjukkan alur kerja sistem yang akan dibangun. Tahap awal dari *flowchart* adalah mengimputkan nilai mahasiswa, selanjutnya menentukan *indiscernibility relation* setiap atribut, jika semua atribut telah ditentukan *indiscernibility relation* maka akan dimasukkan dalam satu himpunan yang bernama *equivalence class*, sehingga setiap atribut memiliki *equivalence class* sendiri. Selanjutnya menghitung *lower* dan *upper approximation* dari satu atribut terhadap atribut lainnya. Hasil dari *lower* dan *upper approximation* dihitung *roughness*-nya. Selanjutnya menghitung nilai rata-rata dari *roughness*-nya dan ditampilkan dalam bentuk tabel. Hasil yang ditampilkan akan dicari nilai *mean roughness* maksimumnya dari setiap atribut. Tahap berikutnya mencari nilai *mean rouhness* maksimum dari semua atribut, atribut yang memiliki nilai *mean roughness* maksimum maka atribut tersebut yang akan membentuk kluster-kluster. Tahap terakhir jika ingin mengulang maka akan diulang ketahap awal dan jika tidak maka proses selesai.

1. Use Case



***Gambar 2*** *Use Case Diagram*

Pada gambar 2 merupakan *use case diagram* memerlukan *actor* sebagai pelaku yang menjalankan case–case yang ada. *Actor* yang terlibat dalam sistem ini secara khusus adalah pengguna yang akan melakukan proses *clustering*. *Actor* dapat mengakses tugas yaitu mengimputkan nilai mahasiswa, melihat nilai, dan melakukan proses *clustering*.

1. Activity Diagram



***Gambar 3*** *Activity Diagram*

Gambar 3 menjelaskan bahwa ketika pertama kali pengguna menjalankan program, pengguna akan masuk ke dalam halaman utama. Setelah masuk pada halaman pertama, pengguna dapat langsung melihat nilai jika sudah pernah diinputkan. Selanjutnya pengguna bisa melakukan *clustering*, jika belum ada nilai maka pengguna harus mengimputkan nilai terlebih dahulu. Setelah data di *input*, proses *clustering* dapat dilakukan sesuai dengan nilai yang terdapat dalam *database*. Setelah proses *clustering* selesai maka akan ditampilkan hasil dalam bentuk tabel dan hasil atribut yang memiliki nilai *mean roughness* maksimum akan ditampilkan, disertai dengan hasil kluster-kluster yang terbentuk.

1. Squence Diagram



***Gambar 4*** *Squence Diagram*

Gambar 4 adalah alur pada *sequence diagram* sama seperti *activity diagram*, mulai dari halaman utama, input data nilai sampai pada hasil proses *clustering*. Hanya saja pada gambar disusun dalam suatu urutan waktu. Garis vertikal merupakan waktu dan garis horizontal merupakan kegiatan yang akan dikerjakan.

### 3.3. Pembuatan Program

Program yang akan diimplementasikan menggunakan bahasa *scripting* PHP, CSS, HTML dan MySQL untuk penyimpanan databasenya. Proses dalam pembuatan program menggunakan perangkat keras dan perangkat lunak sebagai bahan pendukung.

### 3.4. Implementasi Program

Pada tahap implementasi program *clustering* nilai mahasiswa menggunakan metode *Max-Max Roughness* (MMR) dilakukan dengan cara mengunggah aplikasi ke penyedia layanan *hosting*. Kemudian pengguna dapat mengakses program berbasis *web* yang telah diunggah ke *internet* untuk dapat digunakan dalam *clustering* nilai mahasiswa.

**4. IMPLEMENTASI SISTEM**

### 1. Halaman Utama



***Gambar 5*** *Tampilan Halaman Utama*

Pada gambar 5 merupakan tampilan awal saat membuka aplikasi web. Halaman ini berfungsi sebagai tempat untuk menampilkan informasi dari aplikasi web yang sedang dibuka. Halaman utama terdiri dari judul dan satu *button*. Fungsi tombol *button* untuk dapat masuk kehalaman nilai.

### 2. Halaman Nilai Mahasiswa



***Gambar 6*** *Tampilan Halaman Nilai Mahasiswa*

Pada gambar 6 merupakan tampilan halaman nilai mahasiwa. Halaman nilai berfungsi sebagai tempat untuk menampilkan data nilai mahasiswa. Pada halaman nilai akan ditampilkan nilai mahasiswa yang sudah pernah diinputkan sebelumnya, jika belum pernah maka halaman nilai akan kosong dan proses *clustering* nilai mahasiswa tidak dapat dilakukan. Untuk dapat melakukan proses *clustering*, terlebih dahulu menginputkan nilai mahasiswa. Nilai mahasiswa dapat diinputkan dengan dua cara yaitu pertama dengan mengklik *button* input, dengan cara ini data akan diinputkan satu persatu di halaman web atau mengkilik menu import agar dapat memasukkan nilai dalam jumlah banyak, dengan format *file* excel. Pada halaman nilai, juga dapat meng-*export* nilai yang tersimpan didatabase agar dapat digunakan lagi.

### 3. Halaman Import



***Gambar 7*** *Tampilan Halaman Import*

Pada gambar 7 merupakan tampilan halaman *import* nilai mahasiwa. Pada halaman ini terdapat dua *button* yaitu *button* download format dan preview. *Button* download format berfungsi untuk menyesuaikan posisi nilai pada *file* excel untuk dapat dimasukkan kedalam web. Jika nilai yang ingin diinputkan sudah sesuai dengan *format* excel maka dapat mengklik tombol browse dan mencari lokasi *file*. Kilik *button* preview untuk dapat melihat nilai dalam *file* excel, jika nilai sudah benar maka dapat dilakukan *import* data, sebelum mengkilik *button* import terdapat checkbox “hapus data sebelumnya” jika ingin menghapus data nilai yang sudah ada dalam database maka dapat memberi centrang, namun jika ingin menambah nilai yang sudah ada dalam database maka tidak perlu memberi mencetrang.

### 4. Halaman Clustering



***Gambar 8*** *Tampilan Halaman Clustering*

Pada gambar 8 merupakan tampilan halaman *clustering*. Pada halaman ini hasil *Roughness* dan *Mean Roughness* setiap atribut ditampilkan dalam bentuk tabel. Pada halaman ini juga ditampilkan hasil *clustering* nilai mahasiswa.

**5. PENUTUP**

**5.1. Kesimpulan**

Berdasarkan hasil penelitian dari tugas akhir ini, diperoleh kesimpulan bahwa Implementasi Clustering Nilai Mahasiswa Menggunakan Metode Max-Max Roughness dapat digunakan dalam pengelompokkan data. Data yang digunakan dalam proses *clustering* adalah data yang diperoleh dari mahasiswa Teknik Informatika Universitas Teknologi Yogyakarta. Hasil menunjukkan bahwa atribut yang dominan pada atribut lain dapat ditentukan dengan tingkat akurasi 70%. Atribut yang dominan akan menampilkan data dalam satu klaster yang memiliki tingkat kemiripan yang maksimum dan data antar kluster memiliki kemiripan yang minimum. Selain itu metode ini membutuhkan spesifikasi hardware yang cukup untuk dapat melakukan *clustering* dengan jumlah data yang besar, karena melakukan proses perulangan yang cukup banyak untuk membentuk kluster.

**5.2. Saran**

Berdasarkan hasil pembahasan dan kemsimpulan, maka dapat diberikan saran yaitu pengembangan perangkat lunak dapat dikembangkan dengan atribut yang dinamis sehingga jumlah atribut yang digunakan dapat ditentukan dan pengembangan metode dan perangkat lunak perlu dikembangkan dengan lebih baik. Dengan demikian dapat menangani data yang jumlahnya lebih besar dan kompleks.

# **DAFTAR PUSTAKA**

[1] Arief, M. R., Siahaan, D.O. dan Arieshanti, I., (2010), *Klasterisasi Teks Menggunakan Metode Max-Max Roughness(MMR) Dengan Pengayaan Similaritas Kata*, Kursor, Vol *5*(4), 246–255.

[2] Arief, N, T., (2012), *Analisis Implementasi Algoritma Min-Min Roughness ( MMR ) Berbasis Rough Set Theory Dalam Clustering Pada Data Kategorikal Program Studi Sarjana Teknik Informatika Fakultas Informatika Institut Teknologi Telkom Bandung Lembar Pengesahan Analisis Implement*, Tugas Akhir, Universitas Telkom.

[3] Fayyad, U., Shapiro, G.P., dan Smyth, P., (1996), *From Data Mining to Knowledge Discovery in Database, American Association for Artificial Intelligence*, Al Magazine, Vol 17(3), 37-54.

[4] Herawan, T., Ghazali, R., Yanto, I.T.R., dan Deris, M.M., (2010), *Rough Set Approach for Categorical Data Clustering*, International Journal of Database Theory and Application, Vol. 3(1) 33-52.

[5] Irwansyah, E., dan Faisal, M., (2015), *Advanced Clustering Teori dan Aplikasi*, Yogyakarta: Deepublish.

[6] Listiana, N., Anggraeni, W., dan Mukhlason, A., (2011), *Implementasi Algoritma Rough Set Untuk Deteksi dan Penanganan Dini Penyakit Sapi*, Skirpsi, (https://www.google.com/url?q=http://digilib.its.ac.id/public/ITS-Undergraduate-16046-5207100082-Paper.pdf&sa=U&ved=0ahUKEwi49pzX6ufSAhVKp48KHYUkCo4QFggEMAA&client=internal-uds-cse&usg=AFQjCNEz5zcoV6ujvg0Gsj6x93T6G7Nrjw), akses 21 Maret 2017.

[7] Mazlack L, He A., Zhu, Y., and Coppock, S., (2000), *A Rough Set Approach in Choosing Parttion Attributes*, Proceedings of the ISCA 13th International Conference (CAINE-2000), 1-6.

[8] Parmar, D., Wu, T., and Blackhurst, J., (2007), *MMR : An algorithm for clustering categorical data using Rough Set Theory*, Elsevier, Vol 63(3) 879-893.

[9] Pawlak, Z., (1982), *Rough Sets*, International Journal of Computer and Information Sciences, Vol 11(5) 341-356.

[10] Suhendar, A., dan Gunadi, H., (2002), *Visual Modeling Menggunakan UML dan Relational Rose*, Yogyakarta: Infromatika.

[11] Suhirman., (2017), *An Application Of Rough Set Theory To Cluster Student Assessment At Universities*, Jiko, Vol 2(1), 1-9.

[12] Suratiningsih, (2013), *Min-Min Roughness (MMR) untuk Klasterisasi Data Kategori Studi Kasus Kecemasan Belajar pada Mahasiswa Tahun Pertama dan Keempat di UAD*, Konvergensi, Vol 3(2), 19–32.

[13] Tan, P.N., Stenbach., dan Kumar, V., 2006, *introduction to Data Mining*, Pearson Education, Boston.