

## **DATA MINING REKOMENDASI CALON MAHASISWA BERPRESTASI DI STMIK DENPASAR MENGGUNAKAN METODE *TECHNIQUE FOR OTHERS REFERENCE BY SIMILARITY TO IDEAL SOLUTION***

<sup>1)</sup> I Kadek Juni Arta, <sup>2)</sup> Gede Indrawan, <sup>3)</sup> Gede Rasben Dantes

<sup>1,2,3)</sup>Program Studi Ilmu Komputer, Program Pascasarjana  
Universitas Pendidikan Ganesha  
Singaraja, Indonesia

e-mail: [juni.smkdenpasar@gmail.com](mailto:juni.smkdenpasar@gmail.com), [gede.indrawan@gmail.com](mailto:gede.indrawan@gmail.com),  
[rasben.dantes@undiksha.ac.id](mailto:rasben.dantes@undiksha.ac.id)

### **Abstrak**

Perkembangan teknologi yang semakin pesat dan tidak bisa dipungkiri bahwa teknologi sebagai kebutuhan sekunder bagi para penggunanya. Sejauh ini teknologi juga telah diterapkan pada beberapa bidang, salah satu contohnya kehadiran teknologi di sektor pendidikan yang memudahkan para penggunanya untuk melakukan proses penilaian prestasi. Sistem penilaian dan pencatatan yang dilakukan STMIK Denpasar masih konvensional sehingga mempersulit para dosen dan ketua yayasan untuk melakukan analisis lebih lanjut mengenai keadaan calon mahasiswanya. Selain dari itu, penilaian yang dilakukan pun lebih berfokus ke dalam bidang akademis (penguasaan konsep). Permasalahan yang dihadapi oleh pihak manajerial STMIK Denpasar adalah bagaimana memanfaatkan teknologi data mining untuk menentukan prestasi calon mahasiswa baru. Dengan sistem rekomendasi perangkat lunak menyediakan saran untuk *items* yang bisa digunakan oleh user. TOPSIS adalah salah satu metode pengambilan keputusan multikriteria. Metode ini banyak digunakan untuk menyelesaikan pengambilan keputusan secara praktis. Aplikasi telah berhasil memilih mahasiswa berprestasi dengan menggunakan TOPSIS dengan rata-rata access time **0.27268** detik dengan bobot 6 data sampel uji. Implementasi metode TOPSIS pada studi kasus mahasiswa berprestasi dengan menyaring bobot standar minimal hasil TOPSIS untuk diambil 2 nilai rangking terbaik pada ambang batas di atas 0.5. Kinerja TOPSIS dibandingkan dengan metode K-Means Clustering dari aspek access time TOPSIS lebih cepat yaitu **0.27268** detik sedangkan K-Means Clustering **0.3818** detik. Kinerja TOPSIS dibandingkan dengan metode K-Means Clustering dari aspek hasil TOPSIS menghasilkan rekomendasi yang sama dengan K-Means Clustering. Dalam penelitian ini metode TOPSIS baru bisa menghasilkan 2 rekomendasi sehingga pada penelitian selanjutnya metode ini bisa dibandingkan dengan metode lain.

**Kata kunci:** TOPSIS, Rekomendasi, Prestasi, Datamining

### **Abstract**

*Rapid technological developments and can not be denied that technology as a secondary requirement for its users. So far the technology has also been applied in several fields, one example the presence of technology in the education sector that allow its users to make the process of achievement assessment. The assessment system and records STMIK Denpasar still conventional so difficult for lecturers and chairman of the foundation to do further analysis about the state of student candidates. Apart from that, the assessment was more focused on the academic field (mastery of concepts). Problems faced by the managerial STMIK Denpasar is how to leverage data mining technology to determine the achievement of new students. With system software recommendations provide suggestions for items that can be used by the user. TOPSIS is one method of decision-making multicriteria. This method is widely used to complete the decision making in practice. The application has been successfully choose the best students by using TOPSIS with an average access*

*time 0.27268 seconds with 6 weights the data of the test sample. Implementation of TOPSIS method in a case study the best students with a minimum standard weighting filter results TOPSIS to take second best rank value on the threshold above 0.5. TOPSIS performance compared to K-Means Clustering method of aspect TOPSIS faster access time is 0.27268 seconds while K-Means Clustering 0.3818 sec. TOPSIS performance compared to K-Means Clustering method of aspect results of TOPSIS produce the same recommendation to the K-Means Clustering. In this study a new TOPSIS method can generate the two recommendations so that in future studies of this method can be compared with other methods.*

**Keywords : TOPSIS, Recommendations, Achievements, Datamining**

## I. PENDAHULUAN

Perkembangan teknologi yang semakin pesat dan tidak bisa dipungkiri bahwa teknologi sebagai kebutuhan sekunder bagi para penggunanya. Sejauh ini teknologi juga telah diterapkan pada beberapa bidang, salah satu contohnya kehadiran teknologi di sektor pendidikan yang memudahkan para penggunanya untuk melakukan proses penilaian prestasi . Dari perkembangan teknologi tersebut yaitu penilaian prestasi calon mahasiswa pada STMIK Denpasar.

Sistem penilaian dan pencatatan yang dilakukan STMIK Denpasar masih konvensional sehingga mempersulit para dosen dan ketua yayasan untuk melakukan analisis lebih lanjut mengenai keadaan calon mahasiswanya. Selain dari itu, penilaian yang dilakukan pun lebih berfokus ke dalam bidang akademis (penguasaan konsep). Penilaian pengembangan diri dilakukan dengan standar penilaian yang minimum, dan hanya berdasarkan penilaian subjektif dari dosen yang bersangkutan. Sistem penilaian seperti ini tentu saja belum cukup, karena tidak sesuai dengan misi STMIK Denpasar, sekolah ini ingin menghasilkan calon mahasiswa yang bukan hanya pandai di bidang akademis saja, tetapi juga memiliki moral yang baik.

Salah satu alternatif yang dapat dilakukan untuk membantu pihak manajerial meningkatkan kualitas STMIK Denpasar adalah dengan memanfaatkan perkembangan teknologi informasi yang ada, khususnya teknologi sistem data mining. Sistem data mining ini diharapkan dapat memberikan masukan bagi para dosen untuk mengetahui karakter para calon mahasiswanya berdasarkan data yang tersimpan, penilaian keberhasilan proses mengajar, serta membantu pihak panitia seleksi menetapkan langkah/keputusan yang akan diambil berikutnya berdasarkan

analisis data yang ada. Sistem ini juga diharapkan dapat mengatasi beberapa masalah yang terjadi karena penyimpanan berkas data, seperti risiko kehilangan yang besar dan perlunya tempat untuk menyimpan data tersebut, karena sistem ini juga dapat digunakan untuk menyimpan data calon mahasiswa ada dari tahun ke tahun.

Penelitian ini bertujuan untuk merancang sistem data mining dalam penilaian prestasi calon mahasiswa STMIK Denpasar dengan menggunakan metode TOPSIS (*Technique For Others Reference by Similarity to Ideal Solution*) Atribut penilaian yang meliputi nilai raport, tes tulis, kepribadian (wawancara), keaktifan, pekerjaan orang tua, sumber biaya hidup, pekerjaan calon mahasiswa, prestasi semasa sekolah, beasiswa dan kegiatan ekstrakurikuler selama di sekolah.

Permasalahan pada penelitian ini adalah

1. Bagaimana implementasi metode TOPSIS untuk merekomendasi calon mahasiswa berprestasi di STMIK Denpasar ?
2. Bagaimana Kinerja Metode TOPSIS dibandingkan K-Means dalam menentukan Prestasi calon mahasiswa ?

Tujuan penelitian ini adalah

1. Merancang rekomendasi calon mahasiswa berprestasi pada STMIK Denpasar dengan menggunakan metode TOPSIS
2. Membandingkan kinerja metode TOPSIS dan K-Means Clustering yang efektif dalam menentukan calon mahasiswa berprestasi

## II. TINJAUAN PUSTAKA

### 2.1 Data Mining

*Data mining* (Connolly dan Begg, 2010) adalah suatu proses ekstraksi atau penggalian data yang belum diketahui sebelumnya, namun dapat dipahami dan berguna dari *database*

yang besar serta digunakan untuk membuat suatu keputusan bisnis yang sangat penting.

*Data mining* biasa juga disebut dengan “Data atau *knowledge discovery*” atau menemukan pola tersembunyi pada data. *Data mining* adalah proses dari menganalisa data dari perspektif yang berbeda dan menyimpulkannya ke dalam informasi yang berguna.

*Data mining* (Han dan Kamber, 2006) didefinisikan sebagai proses mengekstrak atau menambang pengetahuan yang dibutuhkan dari sejumlah data besar. Pada prosesnya *data mining* akan mengekstrak informasi yang berharga dengan cara menganalisis adanya pola-pola ataupun hubungan keterkaitan tertentu dari data-data yang berukuran besar. *Data mining* berkaitan dengan bidang ilmu-ilmu lain, seperti *Database System*, *Data Warehousing*, *Statistic*, *Machine Learning*, *Information Retrieval*, dan Komputasi Tingkat Tinggi. Selain itu *data mining* didukung oleh ilmu lain seperti *Neural Network*, Pengenalan Pola, *Spatial Data Analysis*, *Image Database*, *Signal Processing*.

Beberapa *survey* tentang proses pemodelan dan metodologi menyatakan bahwa, “*Data mining* digunakan sebagai penunjuk, dimana *data mining* menyajikan intisari atas sejarah, deskripsi dan sebagai standar petunjuk mengenai masa depan dari sebuah proses model *data mining* Karakteristik *data mining* sebagai berikut:

- a. *Data mining* berhubungan dengan penemuan sesuatu yang tersembunyi dan pola data tertentu yang tidak diketahui sebelumnya.
- b. *Data mining* biasa menggunakan data yang sangat besar. Biasanya data yang besar digunakan untuk membuat hasil lebih dapat dipercaya.
- c. *Data mining* berguna untuk membuat keputusan kritis.

Berdasarkan beberapa pengertian tersebut dapat ditarik kesimpulan bahwa *Data Mining* adalah suatu teknik menggali informasi berharga yang terpendam atau tersembunyi pada suatu koleksi data (*database*) yang sangat besar sehingga ditemukan suatu pola yang menarik yang sebelumnya tidak diketahui.

Teknik – teknik data mining telah digunakan untuk menemukan pola yang tersembunyi dan memprediksi tren masa depan. Dan keuntungan kompetitif dari data mining

termasuk dengan meningkatnya pendapatan, berkurangnya pengeluaran, dan kemampuan pemasaran yang meningkat.

*Data mining* dibagi menjadi dua kategori utama (Han dan Kamber, 2006) yaitu:

#### A. Prediktif

Tujuan dari tugas prediktif adalah untuk memprediksi nilai dari atribut tertentu berdasarkan pada nilai atribut-atribut lain. Atribut yang diprediksi umumnya dikenal sebagai target atau variable tak bebas, sedangkan atribut-atribut yang digunakan untuk membuat prediksi dikenal sebagai *explanatory* atau *variable* bebas.

#### B. Deskriptif

Tujuan dari tugas deskriptif adalah untuk menurunkan pola-pola (korelasi, *trend*, *cluster*, teritori, dan anomali) yang meringkas hubungan yang pokok dalam data. Tugas *data mining* deskriptif sering merupakan penyelidikan dan seringkali memerlukan teknik *post-processing* untuk validasi dan penjelasan hasil.

Fungsi dari *data mining* juga ada dalam dunia kesehatan, dimana data mining telah digunakan untuk meningkatkan diagnosis dan pengobatan atau lebih mengerti perilaku dari pasien. (Sandra et all, 2009)

*Data mining* juga memiliki beberapa fungsionalitas yaitu *Concept/Class Description: Characterization and Discrimination, Mining Frequent Patterns, Associations, and Correlations, Classification and Prediction, Cluster Analysis, Outlier analysis*, dan *Evolution analysis*. (Han dan Kamber, 2006 : 21 – 27)

Tujuan dari *data mining* (Hoffer dkk, 2007) adalah:

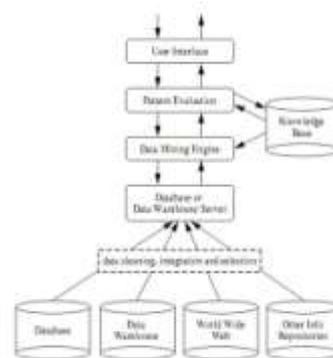
1. *Explanatory*  
Untuk menjelaskan beberapa kondisi penelitian, seperti mengapa penjualan truk *pick-up* meningkat di Colorado.
2. *Confirmatory*  
Untuk mempertegas hipotesis, seperti halnya dua kali pendapatan keluarga lebih suka dipakai untuk membeli peralatan keluarga dibandingkan dengan satu kali pendapatan keluarga.
3. *Exploratory*  
Untuk menganalisa data yang memiliki hubungan yang baru. Misalnya, pola

apa yang cocok untuk kasus penggelapan kartu kredit.

## 2.2 Arsitektur Data Mining

*Data mining* merupakan proses pencarian pengetahuan yang menarik dari data berukuran besar yang disimpan dalam basis data, *data warehouse* atau tempat penyimpanan informasi lainnya. Dengan demikian arsitektur sistem *data mining* memiliki komponen-komponen utama (Han dan Kamber, 2006) yaitu:

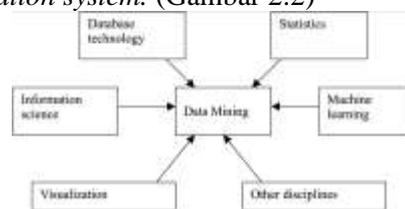
- Database, data warehouse, World Wide Web*, atau tempat penyimpanan informasi lainnya: bisa berbentuk satu atau banyak *database, data warehouse, spreadsheet*, ataupun tempat penyimpanan informasi lainnya. *Data Cleaning, Data Integration* dan *Data Selection* dapat dijalankan pada data tersebut.
- Database dan data warehouse server*. Komponen ini bertanggung jawab dalam pengambilan data yang relevan, berdasarkan permintaan pengguna.
- Knowledge Based*. Komponen ini merupakan domain *knowledge* yang digunakan untuk memandu pencarian atau mengevaluasi pola-pola yang dihasilkan. Pengetahuan tersebut meliputi hirarki konsep yang digunakan untuk mengorganisasikan atribut atau nilai atribut kedalam level abstraksi yang berbeda. Pengetahuan tersebut juga dapat berupa kepercayaan pengguna (*user belief*), yang dapat digunakan untuk menentukan kemenarikan pola yang diperoleh.
- Data mining engine*. Bagian ini merupakan komponen penting dalam arsitektur sistem *data mining*. Komponen ini terdiri dari modul-modul fungsional seperti karakterisasi, asosiasi, klasifikasi, dan analisis *cluster*.
- Graphical user interface (GUI)*. Modul ini berkomunikasi dengan pengguna dan *data mining*. Melalui komponen ini, pengguna berinteraksi dengan sistem menggunakan *query*.



Arsitektur sistem *data mining*

## 2.3 Klasifikasi Sistem Data Mining

*Data Mining* (Han dan Kamber, 2006) merupakan suatu pendekatan dalam pemecahan masalah dengan menggunakan tinjauan berbagai sudut pandang ilmu secara terpadu yaitu, *database system, statistics, machine learning, visualization, dan information system*. (Gambar 2.2)



***Data mining* merupakan irisan dari berbagai disiplin**

## 2.4 Knowledge Discovery In Databases

Han dan Kamber (2006 : 7), lebih spesifik menyatakan istilah *Data Mining* dan *Knowledge Discovery in Databases (KDD)* secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu kumpulan data yang besar. Akan tetapi kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain dan salah satu tahap dalam proses KDD adalah *data mining*.

*Data mining* adalah salah satu langkah dalam proses KDD secara keseluruhan. Secara umum, *data mining* digunakan oleh banyak peneliti sebagai sinonim dari proses KDD. Akhir-akhir ini, *data mining* dan *knowledge discovery* telah diusulkan sebagai nama yang paling memadai untuk keseluruhan proses KDD. *Knowledge Discovery in Databases* berkaitan dengan proses penemuan pengetahuan yang diterapkan pada database. Hal ini juga didefinisikan sebagai proses *non-trivial* untuk identifikasi data yang valid, baru,

berpotensi bermanfaat, dan akhirnya memiliki pola yang dapat dimengerti.

*Knowledge discovery* sering terhalang karena tantangan dalam integrasi dan navigasi dari data yang berbeda. Selain itu, karena jumlah dimensi di dalam data meningkat, pendekatan baru untuk penemuan pola sangat diperlukan. Berdasarkan pengertian beberapa pengertian tersebut dapat ditarik kesimpulan bahwa *Knowledge Discovery in Database* (KDD) adalah proses yang bertujuan untuk menggali dan menganalisis sejumlah besar himpunan data dan mengekstrak informasi serta pengetahuan yang berguna.

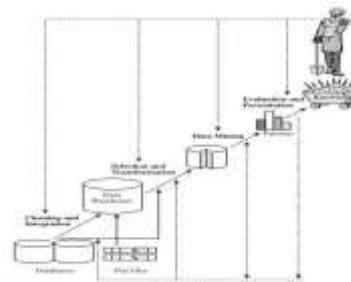
Langkah penting dalam proses KDD dapat dilihat pada gambar 2.3 yang terdiri dari tahapan-tahapan sebagai berikut:

1) *Data cleaning*

*Data cleaning* merupakan proses membuang duplikasi data, memeriksa data yang tidak konsisten, dan memperbaiki kesalahan pada data, seperti kesalahan penulisan. Pada umumnya data yang diperoleh baik dari *database* suatu perusahaan maupun hasil eksperimen, memiliki isi yang tidak sempurna seperti data yang hilang, data yang tidak valid atau juga hanya sekedar salah ketik. Selain itu, ada juga atribut-atribut data yang tidak relevan dengan hipotesa *data mining* yang dimiliki. *Data cleaning* juga akan mempengaruhi hasil informasi dari teknik *data mining* karena data yang ditangani akan berkurang jumlah dan kompleksitasnya.

1) *Data integration*

Proses menambah data yang sudah ada dengan data atau informasi lain yang relevan atau bisa disebut juga merupakan penggabungan data dari berbagai *database* kedalam satu *database* baru yang dibutuhkan oleh KDD. Tahapan *cleaning* dan *integration* pada KDD mengasumsikan bahwa integrator data harus menghapus *noise* dari data awal secara paralel dengan mengintegrasikan beberapa data set.



**Data mining sebagai tahapan dalam proses KDD**

2) *Data selection*

Pemilihan data yang relevan dan dapat dilakukan analisis dari data operasional. Data hasil pemilihan disimpan dalam *database* yang terpisah.

3) *Data transformation*

Proses tranformasi data kedalam bentuk format tertentu sehingga data tersebut sesuai untuk proses *data mining*. Sebagai contoh beberapa metode standar seperti analisis asosiasi dan *clustering* hanya bisa menerima input data kategorikal.

4) *Data mining*

Proses mencari pola atau informasi menarik dengan menggunakan teknik, metode atau algoritma tertentu.

5) *Pattern evaluation*

Mengidentifikasi pola-pola yang benar-benar menarik dari hasil *data mining*. Dalam tahap ini hasil dari teknik *data mining* berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada memang tercapai atau tidak.

6) *Knowledge presentation*

Menampilkan pola informasi yang dihasilkan dari proses *data mining*, visualisasi ini membantu mengkomunikasikan hasil *data mining* dalam bentuk yang mudah dimengerti.

**2.5 TOPSIS**

TOPSIS adalah salah satu metode pengambilan keputusan multikriteria yang pertama kali diperkenalkan oleh Yoon dan Hwang (1981). TOPSIS menggunakan prinsip bahwa alternatif yang terpilih harus mempunyai jarak terdekat dari solusi ideal positif dan terjauh dari solusi ideal negatif dari sudut pandang geometris dengan menggunakan jarak Euclidean untuk menentukan kedekatan relatif dari suatu

alternatif dengan solusi optimal. Solusi ideal positif didefinisikan sebagai jumlah dari seluruh nilai terbaik yang dapat dicapai untuk setiap atribut, sedangkan solusi negatif-ideal terdiri dari seluruh nilai terburuk yang dicapai untuk setiap atribut.

TOPSIS mempertimbangkan keduanya, jarak terhadap solusi ideal positif dan jarak terhadap solusi ideal negatif dengan mengambil kedekatan relatif terhadap solusi ideal positif. Berdasarkan perbandingan terhadap jarak relatifnya, susunan prioritas alternatif bisa dicapai.

Metode ini banyak digunakan untuk menyelesaikan pengambilan keputusan secara praktis. Hal ini disebabkan konsepnya sederhana dan mudah dipahami, komputasinya efisien, dan memiliki kemampuan mengukur kinerja relatif dari alternatif-alternatif keputusan.

### Langkah-langkah Metode TOPSIS

Langkah-langkah yang harus dilakukan untuk memperoleh hasil penilaian dalam metode TOPSIS adalah:

1. Membangun *normalized decision matrix*
  - Elemen  $r_{ij}$  hasil dari normalisasi *decision matrix*  $R$  dengan metode *Euclidean length of a vector* adalah:

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^m x_{ij}^2}} \dots\dots\dots(2.1)$$

2. Membangun *weighted normalized decision matrix*

Dengan bobot  $W = (w_1, w_2, \dots, w_n)$ , maka normalisasi bobot matriks  $V$  adalah:

$$V = \begin{bmatrix} w_1 r_{11} & w_2 r_{12} & \dots & w_n r_{1n} \\ w_1 r_{21} & & & \\ \vdots & & & \\ w_1 r_{m1} & w_2 r_{m2} & \dots & w_n r_{mn} \end{bmatrix} \dots\dots\dots(2.2)$$

3. Menentukan solusi ideal dan solusi ideal negatif.

Solusi ideal dinotasikan  $A^*$ , sedangkan solusi ideal negatif dinotasikan  $A^-$ :

$$A^* = \{ (\max v_{ij} | j \in J), (\min v_{ij} | j \in J), i = 1, 2, 3, \dots, m \} = \{v_{1*}, v_{2*}, \dots, v_{n*}\}$$

$$A^- = \{ (\min v_{ij} | j \in J), (\max v_{ij} | j \in J), i = 1, 2, 3, \dots, m \} = \{v_{1-}, v_{2-}, \dots, v_{n-}\} \dots\dots\dots(2.3)$$

Dimana:

$J = \{j = 1, 2, 3, \dots, n \text{ dan } j \text{ merupakan } \textit{benefit criteria}\}$   
 $J = \{j = 1, 2, 3, \dots, n \text{ dan } j \text{ merupakan } \textit{cost criteria}\}$

4. Menghitung separasi

- $S_i^*$  adalah jarak (dalam pandangan *Euclidean*) alternatif dari solusi ideal didefinisikan sebagai:

$$S_i^* = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^*)^2}, \text{ dengan } i=1,2,3,\dots,m \dots\dots\dots(2.4)$$

- Dan jarak terhadap solusi negatif-ideal didefinisikan sebagai:

$$S_i^- = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^-)^2}, \text{ dengan } i=1,2,3,\dots,m \dots\dots\dots(2.5)$$

5. Menghitung kedekatan relatif terhadap solusi ideal

$$C_i^* = \frac{S_i^-}{S_i^* + S_i^-}, \text{ dengan } 0 < C_i^* < 1 \text{ dan } i=1,2,3,\dots,m \dots\dots\dots(2.6)$$

6. Meranking Alternatif

Alternatif dapat diranking berdasarkan urutan  $C_i^*$ . Maka dari itu, alternatif terbaik adalah salah satu yang berjarak terpendek terhadap solusi ideal dan berjarak terjauh dengan solusi negatif-ideal.

### 2.6 Sistem Rekomendasi

Ada beberapa macam definisi dari sistem rekomendasi. Menurut Mahmood dan Ricci, Sistem Rekomendasi adalah peralatan perangkat lunak dan teknik yang menyediakan saran untuk *items* yang bisa digunakan oleh *user*. Secara general sistem rekomendasi didefinisikan sebagai sistem pendukung yang membantu *user* untuk mencari informasi, produk dan servis (buku, film, music dll) dengan menggabungkan dan menganalisa saran dari *user* lain, yang berarti meninjau dari beberapa pihak dan *user* atribut.

Sistem rekomendasi menjadi sebuah penelitian bidang yang penting sejak munculnya makalah pertama tentang *collaborative-filtering* pada pertengahan 1990an. Tujuan dari sistem rekomendasi adalah menghasilkan rekomendasi yang berguna kepada *user* untuk *items* atau produk yang paling menguntungkan bagi *user*. Sistem Rekomendasi menurut Melville dan Sindhvani, terbagi atas 3 jenis, yaitu:

1. *Content-based Filtering*
2. *Collaborative Filtering*
3. *Hybrid-based Filtering*

### 2.7 K-Means Clustering

*K-Means* termasuk dalam *partitioning clustering* yaitu setiap data harus masuk dalam *cluster* tertentu dan memungkinkan bagi setiap

data yang termasuk dalam *cluster* tertentu pada suatu tahapan proses, pada tahapan berikutnya berpindah ke *cluster* yang lain. K-Means memisahkan data ke *k* daerah bagian yang terpisah, dimana *k* adalah bilangan integer positif. Algoritma K-Means sangat terkenal karena kemudahannya dan kemampuannya untuk mengklasifikasi data besar dan *outlier* dengan sangat cepat (Kusumadewi dkk, 2006)

Berikut adalah langkah-langkah algoritma K-Means :

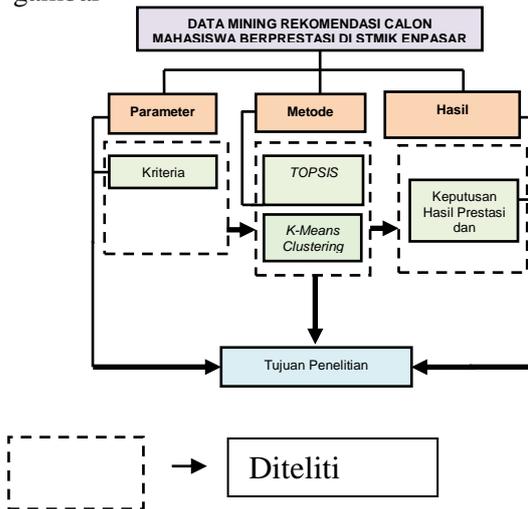
1. Penentuan pusat *cluster* awal  
 Dalam menentukan *n* buah pusat *cluster* awal dilakukan pembangkitan bilangan random yang merepresentasikan urutan data *input*. Pusat awal *cluster* didapatkan dari data sendiri bukan dengan menentukan titik baru, yaitu dengan menrandom pusat awal dari data.
2. Perhitungan jarak dengan pusat *cluster*  
 Untuk mengukur jarak antara data dengan pusat *cluster* digunakan *Euclidian distance*. Algoritma perhitungan jarak data dengan pusat *cluster*
  - Ambil nilai data dan nilai pusat *cluster*
  - Hitung *Euclidian distance* data dengan tiap pusat *cluster*
3. Pengelompokkan data  
 Jarak hasil perhitungan akan dilakukan perbandingan dan dipilih jarak terdekat antara data dengan pusat *cluster*, jarak ini menunjukkan bahwa data tersebut berada dalam satu kelompok dengan pusat *cluster* terdekat. Algoritma pengelompokkan data
  1. Ambil nilai jarak tiap pusat *cluster* dengan data
  2. Cari nilai jarak terkecil
  3. Kelompokkan data dengan pusat *cluster* yang memiliki jarak terkecil.
4. Penentuan pusat *cluster* baru  
 Untuk mendapatkan pusat *cluster* baru bisa dihitung dari rata-rata nilai anggota *cluster* dan pusat *cluster*. Pusat *cluster* yang baru digunakan untuk melakukan iterasi selanjutnya, jika hasil yang didapatkan belum konvergen. Proses iterasi akan berhenti jika telah memenuhi maksimum iterasi yang dimasukkan oleh *User* atau hasil yang dicapai sudah konvergen (pusat *cluster* baru sama dengan pusat *cluster* lama). Algoritma penentuan pusat *cluster* baru
  1. Cari jumlah anggota tiap *cluster*

2. Hitung pusat baru dirumuskan seperti ditunjukkan pada Persamaan (2.7)

$$\text{Pusat baru} = \frac{X_1 + X_2 + X_3 + \dots + X_N}{\text{Jumlah} + 1} \dots \dots \dots (2.7)$$

### III. METODE PENELITIAN

Evaluasi kelayakan mahasiswa merupakan suatu penilaian dimana suatu mahasiswa apakah pantas atau tidak untuk kategori mahasiswa berprestasi. Proses keputusan layak atau tidak dapat dijelaskan dengan gambar



### Hipotesis

Dalam sebuah penelitian, memiliki dugaan sementara mengenai hasil penelitian (hipotesis). Tetapi hipotesis tidak mutlak selalu ada dalam penelitian. Sebelumnya berikut pengertian dari hipotesis. Menurut Sugiyono (2009:159) yang dimaksud dengan hipotesis adalah:

**“Hipotesis diartikan sebagai jawaban sementara terhadap rumusan masalah penelitian.”**

Dari kutipan diatas, hipotesis merupakan suatu pernyataan yang bersifat sementara atau dengan anggapan, pendapat atau asumsi yang mungkin benar dan mungkin salah. Berdasarkan uraian kerangka pemikiran diatas, maka hipotesis yang disajikan penulis adalah **“TOPSIS lebih cepat dalam memproses data pada studi kasus data mining mahasiswa berprestasi di STMIK Denpasar”**.

### METODE

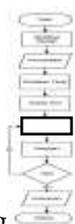
## Desain Penelitian

### Objek yang diteliti

Berdasarkan dari tujuan penelitian, objek yang diteliti adalah data mining rekomendasi calon mahasiswa berprestasi dengan metode *TOPSIS*

### Prosedur Penelitian

Prosedur penelitian yang digunakan dalam dipresentasikan dalam Gambar 3.1:



Gambar 3.1 Kerang lusi Penelitian

### Jenis Penelitian

Pada penelitian ini, peneliti menggunakan metode penelitian kuantitatif. Dimana peneliti mengumpulkan data dan menguji atau membuktikan hipotesis yang ada. Peneliti melakukan survei untuk menentukan frekuensi dan prosentase tanggapan mereka tentang sistem .

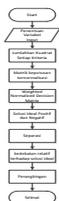
### Sumber Data

Data yang diperoleh dalam penelitian ini ada dua sumber data yaitu:

1. Data Sekunder  
 Penulis akan melakukan pencarian, pembelajaran dari berbagai macam literatur dan dokumen yang menunjang pengerjaan penelitian ini khususnya yang berkaitan dengan *TOPSIS*
2. Data Primer  
 Melakukan pengamatan terhadap data yang diteliti dan melakukan *interview* dengan para pakar yang berkaitan dalam pembuatan aplikasi

### Perancangan TOPSIS

Berikut perancangan TOPSIS dijelaskan pada Gambar 3.3



Gambar 3.2 Flowchart Sistem

## IV HASIL DAN PEMBAHASAN

Hasil pengujian metode topsis dilakukan dengan menggunakan data sampel 6 calon mahasiswa.

Pengujian Hasil TOPSIS menunjukkan 2 mahasiswa yang berprestasi yaitu Calon Mahasiswa 5 dan Calon Mahasiswa 6, dimana Calon Mahasiswa 5 ranking 1 dengan nilai TOPSIS 0.594448 sedangkan Calon Mahasiswa 6 ranking 2 dengan nilai TOPSIS 0.543081.

### Pengujian Access Time Metode TOPSIS

Hasil pengujian access time dilakukan dengan menggunakan 6 sampel data pengujian. Dari 6 data sampel pengujian dengan data sama setiap klik tombol rekomendasi, access time dipengaruhi oleh performance dari komputer atau spesifikasi hardware. Spesifikasi hardware komputer yang digunakan untuk pengujian ini yaitu processor core i7 dengan RAM 16 GB. Selanjutnya, pengujian dilakukan berulang-ulang sebanyak 5 kali klik tombol program rekomendasi. Berikut hasil access time ditunjukkan dalam Tabel 4.1

Tabel 4.1 Pengujian Access Time TOPSIS

No	Pengujian ke	Access Time (detik)
1	1	0.175
2	2	0.4001
3	3	0.3751
4	4	0.2382
5	5	0.175
Rata-Rata (Detik)		<b>0.27268</b>

### Pengujian Metode K-Means Clustering

Hasil pengujian metode K-Means Clustering dilakukan dengan menggunakan data sampel 6 calon mahasiswa.

Pengujian Hasil K-Means Clustering menunjukkan 2 mahasiswa yang berprestasi yaitu Calon Mahasiswa 5 dan Calon Mahasiswa 6.

### Pengujian Access Time Metode K-Means Clustering

Hasil pengujian access time K-means Clustering dilakukan dengan menggunakan 6

sampel data pengujian. Dari 6 data sampel pengujian dengan data sama setiap klik tombol rekomendasi, access time dipengaruhi oleh performance dari komputer atau spesifikasi hardware. Spesifikasi hardware komputer yang digunakan untuk pengujian ini yaitu processor core i7 dengan RAM 16 GB. Selanjutnya, pengujian dilakukan berulang-ulang sebanyak 5 kali klik tombol program rekomendasi. Berikut hasil access time K-means Clustering ditunjukkan dalam Tabel 4.4

Tabel 4.4 Pengujian Access Time K-Means Clustering

No	Pengujian ke	Access Time (detik)
1	1	0.4762
2	2	0.4259
3	3	0.2994
4	4	0.371
5	5	0.3365
Rata-Rata (Detik)		<b>0.3818</b>

### Pengujian Hasil Kedua Metode

Hasil pengujian hasil kedua metode dilakukan dengan menggunakan 6 sampel data pengujian. Dari 6 data sampel pengujian dengan data sama setiap klik tombol rekomendasi, access time dipengaruhi oleh performance dari komputer atau spesifikasi hardware. Spesifikasi hardware komputer yang digunakan untuk pengujian ini yaitu processor core i7 dengan RAM 16 GB. Selanjutnya, pengujian dilakukan berulang-ulang sebanyak 1 kali klik tombol program rekomendasi. Berikut hasil hasil kedua metode ditunjukkan dalam Gambar 4.26



Gambar 4.26 Hasil Perbandingan Hasil Kedua Metode

Gambar 4.26 menunjukkan kedua hasil metode sama-sama valid. Artinya, kedua metode tersebut menunjukkan hasil sama yaitu Calon Mahasiswa 6 dan Calon Mahasiswa 5 sebagai mahasiswa berprestasi.

### Pengujian Access Time Kedua Metode

Hasil pengujian access time kedua metode dilakukan dengan menggunakan 6 sampel data pengujian. Dari 6 data sampel pengujian dengan data sama setiap klik tombol rekomendasi, access time dipengaruhi oleh performance dari komputer atau spesifikasi hardware. Spesifikasi hardware komputer yang

digunakan untuk pengujian ini yaitu processor core i7 dengan RAM 16 GB. Selanjutnya, pengujian dilakukan berulang-ulang sebanyak 1 kali klik tombol program rekomendasi. Berikut hasil access time kedua metode ditunjukkan dalam Gambar 4.29



Gambar 4.29 Hasil Perbandingan Access Time Kedua Metode

Gambar 4.29 menunjukkan kedua access time kedua metode terjadi perbedaan. Yaitu, K-Means 0.3366 sedangkan TOPSIS 0.1834. Dari kedua hasil tersebut dapat disimpulkan TOPSIS mempunyai akses time yang lebih cepat dibandingkan K-Means Clustering



Gambar 2.30 Hasil Perbandingan 2 metode 11 sample

Gambar 4.30 menunjukkan kedua hasil perbandingan access time metode dengan pengujian 11 sampel data. Artinya, kedua metode tersebut menunjukkan hasil berbeda yaitu K-Means membutuhkan waktu lebih banyak yaitu 1,2771 detik sedangkan TOPSIS hanya 0,386 detik seperti pada grafik Gambar 2.31



Gambar 2.31 Grafik Perbandingan Access Time 11 Sample

TOPSIS memiliki akses time lebih cepat dibandingkan dengan K-Means Clustering karena prosedur Topsis sangat mudah, langkah-langkah yang dilakukan untuk memperoleh hasil penilain sudah memiliki tahapan pertahapan tanpa ada pengulangan tahapan. Metode K-Means Clustering kinerjanya lebih lama dibandingkan dengan TOPSIS disebabkan karna penentuan titik centroid yang tidak sesuai akibatnya metode K-Means Clustering akan melakukan beberapa kali iterasi dalam menentukan titik centroid baru sampai proses iterasi berhenti jika sudah

memenuhi maksimum iterasi yang dimasukkan oleh user atau hasil yang dicapai konvergen.

## V. PENUTUP

### Kesimpulan

Dalam memberikan rekomendasi calon mahasiswa berprestasi dengan dengan metode TOPSIS ada prosedur-prosedur yang harus dilakukan, diantaranya adalah : 1) Membuat matriks keputusan yang ternormalisasi. 2) Membuat matriks keputusan yang ternormalisasi terbobot. 3) Menentukan matriks solusi ideal positif dan solusi ideal negative. 4) Menghitung separation measure. 5) Menentukan jarak antara nilai setiap alternatif dengan matriks solusi ideal positif dan negatif. 6) Menentukan nilai preferensi untuk setiap alternatif. 7) Decision matrix D mengacu terhadap m alternatif yang akan dievaluasi berdasarkan n kriteria. 8) Dengan  $x_{ij}$  menyatakan performansi dari perhitungan untuk alternatif ke- $i$  terhadap atribut ke- $j$ .

Kinerja TOPSIS dibandingkan dengan metode K-Means Clustering dari aspek access time TOPSIS lebih cepat yaitu **0.27268** detik sedangkan K-Means Clustering **0.3818** detik.

Kinerja metode TOPSIS yang lebih baik dari metode K-Means Clustering didasari atas prosedur-prosedur yang dilakukan dalam memperoleh hasil penilaian karena dalam metode TOPSIS tidak ada pengulangan prosedur, dibandingkan dengan prosedur dari metode K-Means Clustering yang akan melakukan iterasi secara berulang-ulang apabila dalam proses nya hasil yang didapatkan belum konvergen.

### Saran

Dalam penelitian ini system yang dibuat masih sangat banyak kekurangan, metode TOPSIS sendiri belum mampu memberikan keputusan atau hasil pengelompokan yang lebih dari 2 . Untuk penelitian pengembangan selanjutnya adalah TOPSIS bisa digunakan dalam menentukan keputusan dapat dibandingkan dengan metode yang lain.

## DAFTAR PUSTAKA

Aditiya, Afid. 2016. *Sistem Pendukung Keputusan Klasifikasi Sekolah Berdasarkan Standar Pelayanan*

*Minimal (SPM) dengan Metode K-Means Clustering (Studi Kasus Sekolah Dasar di Lingkungan Dinas Pendidikan Kota Malang)*. Teknik Informatika Institut Teknologi Nasiona ~~K~~Malang

Agustin, Fenty, Eka, M., Fitria, Ardini dan Hanifah, Anif. 2015. Implementasi Algoritma K-Means untuk Menentukan Kelompok Pengayaan Materi Mata Pelajaran Ujian Nasional (Studi Kasus: SMP Negeri 101 Jakarta). *JURNAL TEKNIK INFORMATIKA VOL. 8 NO. 1 APRIL 2015*

Basjir, Mochammad., Supriyanto, Hari dan Suf, Mokh. 2014. *Pengembangan Model Penentuan Prioritas Perbaikan Terhadap Mode Kegagalan Komponen dengan Metodologi FMEA, Fuzzy dan TOPSIS yang terintegrasi*. Institut Teknologi Sepuluh Nopember : Jurusan Teknik Industri

Connoll y, T., Begg, C. 2010. *Database Systems: a practical approach to design, implementation, and management*. 5th Edition. America: Pearson Education.

Fauzan, Achmad., Badharudin, Abid, Yanuar dan Wibowo, Fery. 2014. *Sistem Klasterisasi Menggunakan Metode K-Means dalam Menentukan Posisi Access Point Berdasarkan Posisi Pengguna Hotspot di Universitas Muhammadiyah Purwokerto (Clustering System Using K-Means Method in Determining Access Point Position by Users Hotspot Position at Muhammadiyah University of Purwokerto)*. *JUITA ISSN: 2086-9398 Vol. III Nomor 1, Mei 2014*

Han, Jiawei dan Kamber, Micheline. 2006. *Data Mining : Concept and Techniques Second Edition*. Morgan Kaufmann Publishers.

Hoffer, Jeffrey A., Mary B. Prescott, dan Fred R. McFadden. 2002. *Modern Database Management, 6<sup>th</sup> Edition*. Prentice Hall, Pears on Education Inc, New Jersey.

Kusumadewi, Sri., Hartati, S., Harjoko, A., Wardoyo, R. (2006). *Fuzzy Multi-Attribute Decision Making (FUZZY MADM)*. Yogyakarta: Graha Ilmu.

Lestari, S. 2011. *Seleksi Penerimaan Calon Karyawan Menggunakan Metode TOPSIS*. Konferensi Nasional Sistem dan Informatika 2011, hal: 170-174.

- Muningsih, Elly dan Kiswati, Sri. 2015. Penerapan Metode *K-Means* Untuk *Clustering* Produk *Online Shop* Dalam Penentuan Stok Barang. *Jurnal Bianglala Informatika Vol 3 No 1 Maret 2015*
- Tulus, Sylvia, P dan Hendry. 2014. Perancangan *Clustering Data* Menggunakan Algoritma *K-Means* Berbasis Heatmap (Studi Kasus : Provinsi Papua Barat). *Jurnal Teknologi Informasi-Aiti*, Vol. 11. No.2, Agustus 2014 : 101 – 202
- Yoon, K. dan Hwang, C.L. 1981. *Multiple Attribute Decision Making: Methods and Applications*. Springer Verlag