



# CLUSTERING

**Agung Triayudi, Ph.D., I Wayan Rangga Pinastawa, M.Kom.,  
Jovian Dian Pratama, S.Mat., M.Mat., Sari Ningsih, S.Si., MM.,  
Dr. Ilham, M.Kom., Ari Wibowo, S.E., M.Sc.,  
Loso Judijanto, SSi, MM, MStats., Iqbal Sabilirasyad, S.ST., M.Tr.Kom.,  
Nurhafifah Matondang, S.Kom., M.T.I., Hasnita, M.Si.,  
Dr. Yenny Desnelita, M.Kom., Gustientiedina, M.Kom.**



# CLUSTERING

Agung Triayudi, Ph.D., I Wayan Rangga Pinastawa,  
M.Kom., Jovian Dian Pratama, S.Mat., M.Mat., Sari  
Ningsih, S.Si., MM., Dr. Ilham, S.Kom, M.Kom., Ari Wibowo,  
S.E., M.Sc., Loso Judijanto, SSi., MM., MStats., Iqbal  
Sabilirasyad, S.S.T., M.Tr.Kom., Nurhafifah Matondang,  
S.Kom., M.T.I., Hasnita, M.Si., Dr. Yenny Desnelita, M.Kom.,  
Gustientiedina, M.Kom



## **CLUSTERING**

Copyright© PT Penamudamedia, 2024

### **Penulis:**

Agung Triayudi, Ph.D., I Wayan Rangga Pinastawa, M.Kom., Jovian Dian Pratama, S.Mat., M.Mat., Sari Ningsih.,S.Si.,MM., Dr. Ilham, S.Kom, M.Kom., Ari Wibowo, S.E., M.Sc., Loso Judijanto, SSi., MM., MStats., Iqbal Sabilirrysyad, S.S.T., M.Tr.Kom., Nurhafifah Matondang, S.Kom., M.T.I., Hasnita, M.Si., Dr. Yenny Desnelita, M.Kom., Gustientiedina, M.Kom

### **ISBN:**

978-623-88927-0-9

### **Desain Sampul:**

Tim PT Penamuda Media

### **Tata Letak:**

Enbookdesign

### **Diterbitkan Oleh**

#### **PT Penamuda Media**

Casa Sidoarium RT 03 Ngentak, Sidoarium Dodeam Sleman Yogyakarta

HP/Whatsapp : +6285700592256

Email : penamudamedia@gmail.com

Web : www.penamuda.com

Instagram : @penamudamedia

Cetakan Pertama, Januari 2024

x+187, 15x23 cm

*Hak cipta dilindungi oleh undang-undang*

*Dilarang memperbanyak sebagian atau seluruh isi buku  
tanpa izin Penerbit*



# KATA PENGANTAR

**S**ECARA umum, pengantar buku tentang clustering akan memberikan penjelasan tentang konsep dasar dan prinsip clustering. Clustering adalah teknik dalam analisis data yang bertujuan untuk mengelompokkan objek-objek serupa menjadi kelompok-kelompok yang saling berbeda. Pengantar buku mungkin akan membahas berbagai metode clustering, seperti k-means, hierarchical clustering, dan metode clustering lainnya.

Selain itu, pengantar buku juga mungkin akan menjelaskan tentang aplikasi clustering dalam berbagai bidang, seperti pengelompokan dokumen, analisis citra, analisis biologi, dan lain-lain. Pembaca akan diperkenalkan dengan konsep-konsep penting, seperti ukuran jarak antara objek, evaluasi kualitas clustering, dan teknik pemrosesan data yang relevan.

Pengantar buku clustering juga dapat memberikan gambaran tentang perkembangan terbaru dalam domain ini, seperti clustering dengan metode pembelajaran mesin, penggabungan clustering dengan teknik lain seperti penggalian data, dan aplikasi clustering dalam big data dan analisis data real-time.

Namun, harap diingat bahwa ini adalah deskripsi umum dan dapat bervariasi tergantung pada buku yang spesifik. Jika Anda mencari pengantar khusus dalam buku tertentu, disarankan untuk merujuk pada pengantar yang disediakan di dalam buku tersebut.



# DAFTAR ISI

<b>KATA PENGANTAR .....</b>	<b>v</b>
<b>DAFTAR ISI .....</b>	<b>vii</b>
<b>BAB 1. DEFINISI DAN ISTILAH CLUSTERING .....</b>	<b>1</b>
A. Manfaat Clustering.....	2
B. Konsep dasar Clustering.....	2
C. Jenis – jenis Algoritma Clustering.....	4
D. Cara Kerja Clustering .....	8
E. Pengaplikasian Clustering.....	9
F. Tantangan dan Keterbatasan Clustering.....	10
<b>BAB 2. METODE CLUSTERING .....</b>	<b>12</b>
A. Dasar-dasar Metode Clustering .....	12
B. Jenis-jenis Metode Clustering .....	14
<b>BAB 3. HIERARCHICAL CLUSTERING.....</b>	<b>22</b>
A. Metode Agglomerative .....	24
B. Metode Divisive.....	27
C. Kriteria Penggabungan (Linkage Criteria).....	29
D. Representasi Visual dan Interpretasi Hierarki.....	30

<b>BAB 4. DENSITY-BASED CLUSTERING .....</b>	<b>33</b>
A. Klastering Berbasis Sel kotak ( <i>Grid-Based Clustering</i> : GBC).....	35
B. Kekuatan dan Batasan.....	37
C. Contoh penerapan <i>Density-Based Clustering</i> (DBSCAN).....	38
D. Implementasi <i>Density-Based Clustering</i> .....	39
<b>BAB 5. CLUSTERING IN UNSUPERVISED LEARNING.....</b>	<b>43</b>
A. Jaringan saraf .....	43
B. Jaringan Tertentu .....	49
C. Perbandingan jaringan .....	53
D. Hebbian , SENI, SOM .....	57
E. Metode Probabilistik.....	57
F. Pendekatan.....	58
G. Metode momen .....	59
H. Pengelompokan dalam Pembelajaran Mesin Tanpa Pengawasan	60
I. Analisis pengelompokan.....	67
<b>BAB 6. CLUSTERING IN MACHINE LEARNING .....</b>	<b>77</b>
A. Definisi clustering.....	77
B. Algoritma clustering yang umum digunakan .....	79
C. Metrik evaluasi clustering .....	85
D. Tantangan dalam mengevaluasi hasil clustering.....	86
E. Contoh studi kasus evaluasi clustering.....	88
F. Aplikasi Clustering dalam Machine Learning .....	89



G. Tantangan dan Inovasi Terkini .....	92
<b>BAB 7. DATA SETS DALAM CLUSTERING ANALYSIS .....</b>	<b>95</b>
A. Kesesuaian <i>Data Set</i> dengan Pilihan Algoritma <i>Clustering</i> .....	97
B. Pra-Pengolahan <i>Data Set</i> .....	99
C. Pemilahan Data <i>Training, Testing</i> , dan Validasi .....	101
D. Proporsi Training Data Dan Testing Data .....	102
<b>BAB 8. IMPLEMENTASI CLUSTERING KNOWLEDGE</b>	
<b>KOMBINATION WITH CLUSTERING .....</b>	<b>118</b>
A. Clustering .....	118
B. Algoritma Yang Digunakan Dalam Knowledge Combination With Clustering .....	119
C. Perbedaan Antara Algoritma K-Means Dan Algoritma Prototype Dalam Knowledge Combination With Clustering.....	121
D. Cara Mengaplikasikan Knowledge Combination With Clustering Dalam Analisis Data .....	122
E. Kelebihan Dan Kekurangan Dari Algoritma K-Means Dan Algoritma Prototype Dalam Knowledge Combination With Clustering.....	123
F. Cara Menentukan Parameter Yang Tepat Untuk Algoritma Prototype Dalam Knowledge Combination With Clustering.....	125
G. Parameter Yang Perlu Diperhatikan Dalam Algoritma Prototype Dalam Knowledge Combination With Clustering.....	125
H. Cara Menentukan Parameter Yang Tepat Untuk Algoritma K-Means Dalam Knowledge Combination With Clustering .....	126
I. Teknik Yang Dapat Digunakan Untuk Menentukan Jumlah Kluster Yang Tepat Dalam Algoritma K-Means Dalam Knowledge Combination With Clustering .....	127

J. Terdapat Beberapa Hasil Penelitian Yang Dapat Di Kombinasikan Dengan Clustering .....	128
<b>BAB 9. BIG DATA.....</b>	<b>132</b>
A. Pengenalan Big Data.....	132
B. Sistem Penyimpanan Skala Besar.....	136
C. Teknologi Pengolahan Data.....	136
D. Teknik Analisis Data.....	137
E. Visualisasi Big Data pada metode Clustering .....	140
<b>BAB 10. PARTITIONAL CLUSTERING.....</b>	<b>147</b>
A. MacQueen’s Method.....	153
B. Forgy’s Method .....	153
C. Jancey’s Method.....	154
<b>BAB 11. MODEL BASED CLUSTERING.....</b>	<b>159</b>
A. Model Finite Mixture.....	161
B. Model Finite Mixture Normal Multivariat.....	163
C. Model Finite Mixture t Multivariat.....	164
D. COBWEB : Pendekatan Statistik dan AI .....	165
<b>DAFTAR PUSTAKA .....</b>	<b>172</b>
<b>TENTANG PENULIS.....</b>	<b>182</b>

# CLUSTERING

Clustering adalah teknik dalam analisis data yang digunakan untuk mengelompokkan objek-objek data menjadi kelompok-kelompok yang serupa berdasarkan kesamaan karakteristik atau atribut mereka. Tujuan utama clustering adalah menemukan pola atau struktur tersembunyi dalam data tanpa menggunakan label kelas sebelumnya.

Dalam buku ini, pembaca akan diperkenalkan pada konsep dasar clustering, metode-metode yang umum digunakan, dan algoritma-algoritma yang efektif untuk melakukan clustering. Buku ini juga membahas berbagai masalah dan tantangan yang dihadapi dalam clustering, seperti pemilihan jumlah kelompok yang optimal, menangani data yang tidak lengkap atau noisy, serta memilih metrik jarak yang sesuai.

Pembaca akan mempelajari tentang teknik-teknik clustering yang populer, seperti k-means, hierarchical clustering, DBSCAN, dan masih banyak lagi. Buku ini juga akan membahas strategi evaluasi clustering yang dapat digunakan untuk mengukur kualitas dan keefektifan hasil clustering.

Dengan pemahaman yang mendalam tentang clustering, pembaca akan dapat mengaplikasikan teknik ini dalam berbagai bidang, seperti analisis data bisnis, pengelompokan konsumen, pengenalan pola dalam citra atau teks, dan banyak lagi. Buku ini juga akan memberikan contoh nyata dan studi kasus yang mengilustrasikan penggunaan clustering dalam berbagai konteks.

Penulis buku ini berharap bahwa pembaca akan mendapatkan pemahaman yang kuat tentang clustering dan mampu mengaplikasikan teknik ini dengan baik dalam analisis data mereka. Diharapkan buku ini akan menjadi panduan yang berharga dan bermanfaat bagi para peneliti, praktisi, dan mahasiswa yang tertarik dalam analisis data dan pengelompokan objek data.



PT Penerbit Penamuda Media  
Godean, Yogyakarta  
085700592256  
@penamuda\_media  
penamuda.com