



DATA MINING

▲ +9.803.289

▲ +6.768.223

▲ +5.989.834

Andi Wijaya, M. Kom., I Wayan Rangga Pinastawa, M.Kom.,
Kharisma Wiati Gusti, S.T, M.T., Petrus Christo, M.Kom.,
Loso Judijanto, S.Si, M.M., MStats., Muchamad Kurniawan, S.Kom., M.Kom.,
Muhamad Risqiwahid, S.Kom., Budianingsih, S.T., M.T.,
Iqbal Sabilirasyad, S.ST., M.Tr.Kom., Afifah Nurul Izzati, S.Kom., M.Kom.,
Jovian Dian Pratama, S.Mat., M.Mat., Zuriati, S.Kom., M.Kom.,
Sriyanto, S.Kom., M.M., Ph.D.

▼ -2.694.521

DATA MINING

Andi Wijaya, M.Kom., I Wayan Rangga Pinastawa, M.Kom.,
Kharisma Wiati Gusti, S.T., M.T., Petrus Christo, M.Kom.,
Loso Judijanto, S.Si., M.M., M.Stats., Muchamad Kurniawan, S.Kom., M.Kom.,
Muhamad Risqiwahid, S. Kom., Budianingsih, S.T., M.T.,
Iqbal Sabilirasyad, S.S.T., M.Tr.Kom., Afifah Nurul Izzati, S.Kom., M.Kom.,
Jovian Dian Pratama, S.Mat., M.Mat., Zuriati, S.Kom., M.Kom.,
Sriyanto, M.M., Ph.D

DATA MINING

Copyright© PT Penamudamedia, 2024

Penulis:

Andi Wijaya, M.Kom., I Wayan Rangga Pinastawa, M.Kom.,
Kharisma Wiati Gusti, S.T., M.T., Petrus Christo, M.Kom.,
Loso Judijanto, S.Si., M.M., M.Stats., Muchamad Kurniawan, S.Kom., M.Kom.,
Muhamad Risqiwahid, S. Kom., Budianingsih, S.T., M.T.,
Iqbal Sabilirasyad, S.S.T., M.Tr.Kom., Afifah Nurul Izzati, S.Kom., M.Kom.,
Jovian Dian Pratama, S.Mat., M.Mat., Zuriati, S.Kom., M.Kom., Sriyanto, M.M., Ph.D

ISBN:

978-623-8586-01-1

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, Maret 2024

xii + 181, 15x23 cm

Hak cipta dilindungi oleh undang-undang

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

Kata Pengantar

Hormat kami,

Dengan penuh kebanggaan, kami mempersembahkan buku ini tentang data mining. Buku ini bertujuan untuk membawa pembaca dalam perjalanan yang menarik dan bermanfaat dalam dunia pengeksploasian data yang kaya.

Data mining, sebagai landasan penting dalam analisis data, memberikan wawasan berharga yang dapat mengubah bagaimana kita memandang informasi. Melalui teknik-teknik canggihnya, data mining memungkinkan kita untuk mengidentifikasi pola tersembunyi, meramalkan tren, dan membuat keputusan yang lebih cerdas.

Dalam buku ini, kami menggali konsep-konsep data mining secara komprehensif, mulai dari metode clustering hingga analisis asosiasi, memberikan pemahaman yang kokoh tentang bagaimana data mining dapat diterapkan dalam berbagai konteks, mulai dari bisnis hingga ilmu pengetahuan.

Kami berharap buku ini tidak hanya menjadi panduan praktis bagi para profesional dan peneliti yang tertarik dalam analisis

data, tetapi juga menjadi sumber inspirasi bagi mereka yang ingin mendalami potensi besar yang tersimpan dalam setiap titik data.

Terakhir, kami ingin menyampaikan terima kasih kepada semua individu yang telah memberikan kontribusi dalam pembuatan buku ini. Semoga buku ini tidak hanya memberikan wawasan yang berharga, tetapi juga memperkaya pengetahuan dan pengalaman pembaca dalam dunia yang semakin terhubung melalui data.

Daftar Isi

KATA PENGANTAR.....	v
DAFTAR ISI.....	vii

BAB 1

PENGERTIAN DATA MINING	1
A. Definisi Data Mining.....	3
B. Tujuan dan Manfaat Data Mining	4
C. Teknik Data Mining	4
D. Tantangan dalam Data Mining	6
E. Aplikasi Data Mining	6

BAB 2

SEJARAH DATA MINING.....	9
A. Awal Mula Data Mining.....	10
B. Masa Pertengahan Data Mining	10
C. Inovasi Data Mining Pada Era Modern	13

D. Kontribusi Perusahaan dan Institusi	15
E. Tantangan dan Perkembangan Terkini	16
F. Pandangan Masa Depan Data Mining.....	17

BAB 3

DATA WAREHOUSE.....	19
A. Konsep Dasar Data Warehouse.....	20
B. Model Data Warehouse.....	25
C. Implementasi Data Warehouse untuk Data Mining	27

BAB 4

ARSITEKTUR DATA MINING.....	31
A. Sumber Data	32
B. Pemrosesan Data	32
C. Arsitektur Penambangan Data.....	34
D. Proses Penambangan.....	37
E. Evaluasi dan Interpretasi	38

BAB 5

SISTEM KLASIFIKASI DALAM DATA MINING.....	41
A. Definisi Klasifikasi	42
B. Tujuan dan Manfaat Klasifikasi.....	42
C. Asumsi dan Prasyarat.....	44
D. Kelebihan dan Kekurangan Masing-masing Algoritma.....	51

E. Algoritma Logistic Regression	52
F. Algoritma Naive Bayes.....	53
G. Algoritma Decision Tree.....	55
H. Algoritma Random Forest	56
I. Algoritma K-Nearest Neighbour (KNN).....	57
J. Algoritma Artificial Neural Network (ANN)	59

BAB 6

PRINSIP DATA MINING	61
A. Prinsip Analisis Prediksi	68
B. Prinsip Asosiasi	71
C. Prinsip Pengelompokan	72

BAB 7

SOFTWARE DATA MINING	75
A. Pengertian Data Mining dan <i>Software Data Mining</i>	76
B. Pertimbangan Pemilihan <i>Software</i>	77
C. Kategori <i>Software Data Mining</i>	77
D. Contoh <i>Software Data Mining</i>	78

BAB 8

PEMAHAMAN RAPIDMINER	85
A. Apa itu RapidMiner	86
B. Mengapa RapidMiner	86

C. Instalasi RapidMiner.....	87
D. Memulai Penggunaan RapidMiner	89
E. Cara import data set ke dalam proses.....	91

BAB 9

PENGUJIAN METODE CLUSTERING	95
A. Probabilistic Models for Clustering.....	98
B. Partitional dan Hierarchical Clustering Algorithms.....	99
C. Density-Based Clustering.....	101
D. Grid-Based Clustering	102
E. Nonnegative Matrix Factorizations for Clustering.....	104
F. Spectral Clustering.....	105

BAB 10

ALGORITMA FP-GROWTH.....	107
A. Algoritma <i>FP-Growth</i> Pada Rapidminer	110

BAB 11

ALGORITMA K-MEANS PADA RAPIDMINER.....	119
A. Konsep Dasar Algoritma K-Means	120
B. RapidMiner.....	120
C. K-Means pada RapidMiner.....	122
D. Proses Tutorial	129
E. Implementasi pada RapidMiner	130

F. Prosedur K-Means menggunakan RapidMiner.....	131
---	-----

BAB 12

ALGORITMA NAIVE BAYES PADA RAPIDMINER.....	135
---	------------

A. Teori Algoritma Naïve Bayes.....	136
B. Implementasi Algoritma Naïve Bayes Menggunakan Rapidminer.....	138

BAB 13

ALGORITMA DECISION TREE PADA RAPIDMINER.....	151
---	------------

A. Teori Algoritma Decision Tree.....	152
B. Implementasi Algoritma Decision Tree Menggunakan Rapidminer.....	154

DAFTAR PUSTAKA.....	165
----------------------------	------------

TENTANG PENULIS	174
------------------------------	------------

1

PENGERTIAN DATA MINING

Andi Wijaya, M. Kom.

Di tengah kemajuan era digital yang tak henti, jumlah data yang dihasilkan oleh perusahaan, organisasi, dan individu melesat dengan cepat tiap harinya. Data yang terbentuk mencakup ragam jenis, mulai dari teks, gambar, audio, video, hingga data terstruktur dari basis data konvensional. Bagaimanakah kita dapat mengoptimalkan nilai dari keberadaan data ini? Menjawab pertanyaan tersebut menjadi pokok perhatian utama dalam ranah ilmu data mining.

Dengan laju pertumbuhan yang eksponensial, data menjadi aset penting dalam era digital. Perlu dicatat bahwa sifatnya yang beragam, mencakup format dan jenis yang berbeda-beda, menawarkan tantangan unik dan peluang yang luar biasa dalam pemahaman dan pemanfaatan. Oleh karena itu, menjadi esensial untuk memahami bagaimana data mining dapat membantu dalam menyaring, menganalisis, dan mengekstraksi wawasan berharga dari hiruk-pikuk data ini.

Melalui ilmu data mining, kita dapat mengembangkan teknik-teknik yang memungkinkan identifikasi pola, korelasi, dan informasi tersembunyi yang mungkin tidak dapat dilihat secara langsung. Dengan kemampuan ini, organisasi dapat membuat keputusan yang lebih terinformasi, memprediksi tren masa depan, dan mengoptimalkan proses bisnis mereka.

Dalam konteks ini, data mining menjadi landasan bagi penelitian dan pengembangan berbagai algoritma, metode analisis, dan model prediktif. Kemajuan dalam bidang ini tidak hanya membuka peluang baru dalam berbagai industri, tetapi juga menantang kita untuk terus memperbaiki pendekatan dan teknik yang ada guna menghadapi kompleksitas data yang semakin bertambah.

DATA MINING

Data mining adalah proses mengekstrak pola atau pengetahuan yang berguna dari kumpulan data yang besar. Teknik ini digunakan untuk menganalisis data secara mendalam, mengidentifikasi hubungan yang tersembunyi, dan membuat prediksi berdasarkan informasi yang terungkap. Dengan menggunakan berbagai metode statistik, matematika, dan kecerdasan buatan, data mining membantu organisasi dalam mengambil keputusan yang lebih cerdas dan strategis.

Tujuan utama dari data mining adalah mengubah data mentah menjadi informasi yang berguna dan bermakna. Dengan menerapkan teknik seperti clustering, klasifikasi, regresi, dan asosiasi, data mining dapat mengungkap pola-pola yang membantu dalam pemahaman perilaku pelanggan, analisis pasar, deteksi penipuan, prediksi tren, dan banyak lagi.

Data mining digunakan di berbagai industri seperti perbankan, ritel, kesehatan, dan telekomunikasi untuk meningkatkan efisiensi operasional, memahami kebutuhan pelanggan, dan mengoptimalkan strategi bisnis. Dengan pertumbuhan volume data yang cepat, data mining menjadi semakin penting dalam mengubah data menjadi wawasan berharga yang mendukung pengambilan keputusan yang lebih baik.

ISBN 978-623-8586-01-1



9

786238

586011



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