

**Tim Penulis:**

**Widyastuti Andriyani, Fauzan Natsir, Hendarman Lubis, Sasmi Hidayatul Yulianing Tyas,  
Dwiny Meidelfi, Safira Faizah, Nurlaida, Heri Kurniawan, Irma Wahyuningtyas,  
Firman Noor Hasan, Irfan Ricky Afandi, Sindrawati, Erna Hikmawati**

# ***PERANGKAT LUNAK*** ***DATA MINING***



# ***PERANGKAT LUNAK*** ***DATA MINING***

**Tim Penulis:**

**Widyastuti Andriyani, Fauzan Natsir, Hendarman Lubis, Sasmi Hidayatul Yulianing Tyas,  
Dwiny Meidelfi, Safira Faizah, Nurlaida, Heri Kurniawan, Irma Wahyuningtyas,  
Firman Noor Hasan, Irfan Ricky Afandi, Sindrawati, Erna Hikmawati**

# PERANGKAT LUNAK *DATA MINING*

Tim Penulis:

**Widyastuti Andriyani, Fauzan Natsir, Hendarman Lubis, Sasmi Hidayatul Yulianing Tyas, Dwiny Meidelfi, Safira Faizah, Nurlaida, Heri Kurniawan, Irma Wahyuningtyas, Firman Noor Hasan, Irfan Ricky Afandi, Sindrawati, Erna Hikmawati**

Desain Cover:

**Septian Maulana**

Sumber Ilustrasi:

**www.freepik.com**

Tata Letak:

**Handarini Rohana**

Editor:

**Dr. Arief Andy Soebroto, S.T., M.Kom.**

**Yudi Ramdhani, S.T., M.Kom.**

ISBN:

**978-623-500-504-1**

Cetakan Pertama:

**November, 2024**

---

Hak Cipta Dilindungi Oleh Undang-Undang

**by Penerbit Widina Media Utama**

---

Dilarang keras menerjemahkan, memfotokopi, atau memperbanyak sebagian atau seluruh isi buku ini tanpa izin tertulis dari Penerbit.

**PENERBIT:**

**WIDINA MEDIA UTAMA**

Komplek Puri Melia Asri Blok C3 No. 17 Desa Bojong Emas  
Kec. Solokan Jeruk Kabupaten Bandung, Provinsi Jawa Barat

**Anggota IKAPI No. 360/JBA/2020**

Website: [www.penerbitwidina.com](http://www.penerbitwidina.com)

Instagram: @penerbitwidina

Telepon (022) 87355370

# KATA PENGANTAR

Puji syukur kepada Tuhan yang penuh Kasih, sehingga buku "Perangkat Lunak *Data Mining*" selesai kami susun. Sebagai salah satu penulis, saya merasa terhormat mempersembahkan karya ini yang melibatkan kolaborasi dari para ahli di bidang *Data Mining*, analisis data, dan pengembangan perangkat lunak. Proses penulisan buku ini merupakan perjalanan dan penuh inspirasi, di mana kami sebagai tim berupaya untuk menyatukan berbagai perspektif dan pengalaman kami dalam dunia teknologi, khususnya terkait dengan pemanfaatan perangkat lunak dalam pengelolaan data yang semakin kompleks.

Dalam buku ini, kami membahas peran sentral perangkat lunak dalam *Data Mining*, yang memungkinkan pengguna untuk mengekstrak wawasan berharga dari data besar dan tidak terstruktur. Kolaborasi ini bukan hanya sekadar pertukaran pengetahuan, tetapi juga upaya untuk menjelaskan bagaimana berbagai algoritma dan teknik yang ada dapat diterapkan secara praktis dengan bantuan perangkat lunak canggih. Dengan pendekatan ini, kami berharap buku ini dapat menjadi panduan bagi para pembaca yang tertarik untuk memahami lebih jauh dunia *Data Mining*, mulai dari dasar hingga teknik yang lebih kompleks seperti *Decision Trees*, *Random Forest*, dan *Support Vector Machines (SVM)*.

Kami juga ingin menyampaikan terima kasih yang tulus kepada semua pihak yang telah mendukung proses penulisan buku ini, termasuk kolega di berbagai lembaga akademik dan industri yang telah memberikan masukan berharga, serta para editor yang telah berkontribusi dalam meningkatkan kualitas buku ini. Harapan kami, buku "Perangkat Lunak *Data Mining*" ini tidak hanya memberikan pemahaman tentang peran perangkat lunak dalam analisis data, tetapi juga menginspirasi pembaca untuk terus mengeksplorasi dan berinovasi dalam bidang ini yang terus berkembang pesat. Selamat membaca, semoga menemukan wawasan baru dan berguna dari buku ini.

Yogyakarta, Oktober 2024

**Penulis**

# DAFTAR ISI

<b>KATA PENGANTAR</b> .....	<b>iii</b>
<b>DAFTAR ISI</b> .....	<b>iv</b>
<b>BAB 1 DEFINISI DATA MINING DAN PERANGKAT LUNAK</b> .....	<b>1</b>
A. Latar Belakang .....	1
B. Definisi <i>Data Mining</i> .....	3
C. Kategori Perangkat Lunak <i>Data Mining</i> .....	6
D. Kriteria Pemilihan Perangkat Lunak <i>Data Mining</i> .....	9
E. Kriteria Pemilihan Perangkat Lunak <i>Data Mining</i> .....	10
F. Kesimpulan .....	11
<b>BAB 2 PENGUMPULAN DATA DAN PRA-PEMROSESAN DATA</b> .....	<b>13</b>
A. Pengantar Pengumpulan Data ( <i>Data Collection</i> ) .....	13
B. Pengantar Pra-Pemrosesan Data ( <i>Data Preprocessing</i> ) .....	15
C. Tahapan Pra-Pemrosesan Data ( <i>Data Preprocessing</i> ) .....	18
D. Penutup .....	20
<b>BAB 3 ARSITEKTUR PERANGKAT LUNAK</b> .....	<b>23</b>
A. Komponen Utama .....	23
B. Antarmuka Pengguna (UI) .....	24
C. Modul Pemrosesan Data .....	26
D. Modul Analisis dan Algoritma .....	27
E. Modul Penyimpanan Data .....	29
F. Integrasi dengan Sistem Lain .....	30
G. <i>Database</i> .....	32
H. Api Eksternal .....	33
<b>BAB 4 DESICION TREE</b> .....	<b>37</b>
A. Pengantar <i>Decision Tree</i> .....	37
B. Struktur <i>Decision Tree</i> .....	37
C. Komponen <i>Decision Tree</i> .....	39
D. Algoritma Pembuatan <i>Decision Tree</i> .....	39
E. Evaluasi Kinerja <i>Decision Tree</i> .....	44
F. Implementasi Perangkat Lunak untuk <i>Decision Tree</i> .....	44
G. Studi Kasus <i>Decision Tree</i> .....	48

<b>BAB 5 RANDOM FOREST</b> .....	<b>51</b>
A. Cara Kerja <i>Random Forest</i> .....	51
B. Aplikasi <i>Random Forest</i> .....	52
C. Perbandingan dengan Algoritma Lain.....	53
D. Implementasi dalam <i>Python</i> .....	54
E. Perbandingan <i>Random Forest</i> dengan Algoritma <i>Ensemble</i> Lainnya.....	56
F. Perbandingan Kinerja <i>Random Forest</i> dengan Algoritma <i>Deep Learning</i> .....	57
G. Teknik Visualisasi Hasil Prediksi <i>Random Forest</i> .....	59
<b>BAB 6 SUPPORT VECTOR MACHINE (SVM)</b> .....	<b>61</b>
A. Pengantar .....	61
B. Klasifikasi SVM.....	62
C. Pembelajaran dan Generalisasi.....	67
D. Klasifikasi <i>Soft Margin</i> .....	68
E. Trik Kernel.....	68
F. Metode Sequential.....	69
G. Implementasi SVM .....	69
<b>BAB 7 K-MEANS</b> .....	<b>71</b>
A. <i>Clustering</i> .....	71
B. <i>K-Means</i> .....	73
C. Aplikasi <i>K-Means</i> pada <i>Python</i> .....	78
D. Pemilihan Jumlah <i>Cluster</i> (K).....	80
E. Kelemahan <i>K-Means</i> .....	81
<b>BAB 8 ANALISIS KLASTER HIERARKI</b> .....	<b>85</b>
A. Masalah Penelitian untuk Analisis Kluster .....	85
B. Jenis Data, Asumsi dan Sampling .....	86
C. Prosedur Analisis Kluster dan Interpretasi.....	87
D. Praktik Analisis Kluster dengan <i>Software</i> .....	92
<b>BAB 9 APRIORI ALGORITHM</b> .....	<b>99</b>
A. Pengantar .....	99
B. Pengertian Algoritma Apriori.....	99
C. Cara Kerja Algoritma Apriori .....	100
D. Penerapan Algoritma Apriori .....	101
E. Kelebihan dan Kelemahan Algoritma Apriori .....	103
F. Contoh Studi Kasus.....	104

G. Penutup .....	108
<b>BAB 10 ALGORITMA FP-GROWTH .....</b>	<b>109</b>
A. Pengantar .....	109
B. Definisi .....	110
C. <i>Frequently Pattern Tree (FP-Tree)</i> .....	110
D. Cara Kerja Algoritma <i>FP-Growth</i> .....	113
E. Kelebihan dan Kekurangan dari Algoritma <i>FP-Growth</i> .....	116
F. Apriori vs <i>FP-Growth</i> .....	117
G. Kesimpulan .....	118
<b>BAB 11 TANTANGAN DALAM PENGEMBANGAN PERANGKAT LUNAK</b>	
<b><i>DATA MINING</i> .....</b>	<b>119</b>
A. Pengantar .....	119
B. Tantangan Perangkat Lunak <i>Data Mining</i> .....	120
C. Studi Kasus Tantangan Pengembangan Perangkat Lunak <i>Data Mining</i> di Industri Ritel <i>e-Commerce</i> .....	126
D. Penutup .....	128
<b>BAB 12 PERKEMBANGAN PERANGKAT LUNAK DIMASA DEPAN .....</b>	<b>131</b>
A. Pengantar .....	131
B. Definisi, Peran dan Pentingnya Perangkat Lunak .....	132
C. Tren Teknologi Terkini yang Mempengaruhi Perkembangan Perangkat Lunak .....	133
D. Evolusi Metodologi Pengembangan Perangkat Lunak.....	135
E. Perkembangan dalam Pengalaman Pengguna ( <i>User Experience</i> ).....	135
F. Peran Keamanan dan Privasi dalam Pengembangan Perangkat Lunak .....	136
G. Penggunaan Perangkat Lunak dalam Industri Spesifik .....	136
H. Tantangan yang Akan Dihadapi dalam Pengembangan Perangkat Lunak.....	136
I. Masa Depan Pengembangan Perangkat Lunak .....	137
J. Kesimpulan.....	138
<b>BAB 13 IMPLEMENTASI <i>DATA MINING</i> DENGAN PERANGKAT LUNAK</b>	
<b><i>ORANGE</i>.....</b>	<b>141</b>
A. Pendahuluan.....	141
B. Implementasi Algoritma <i>Decision Tree</i> pada <i>Orange</i> .....	143
<b>DAFTAR PUSTAKA .....</b>	<b>152</b>
<b>PROFIL PENULIS .....</b>	<b>164</b>

# PERANGKAT LUNAK DATA MINING

Buku "Perangkat Lunak *Data Mining*" menyajikan teknik dan alat yang digunakan dalam penggalian data untuk membantu pengambil keputusan memahami serta memanfaatkan data dalam cara yang paling efektif. Dengan struktur ini, dapat menambah wawasan tentang bagaimana perangkat lunak *data mining* beroperasi dan bagaimana memaksimalkannya untuk berbagai aplikasi analitis.

Buku ini menjelaskan konsep dasar *data mining* serta pentingnya perangkat lunak dalam proses pengumpulan dan pra-pemrosesan data. Pembaca akan diajak untuk memahami arsitektur perangkat lunak yang umum digunakan, termasuk berbagai metode dan algoritma yang menjadi andalan, seperti *Decision Trees*, *Random Forest*, *Support Vector Machines (SVM)*, *K-Means*, dan *Hierarchical Clustering*. Selain itu, buku ini mengupas dua algoritma populer dalam analisis data, yaitu *Apriori* dan *FP-Growth*, yang membantu dalam menemukan pola dalam data besar.

Buku ini juga membahas tantangan yang sering dihadapi dalam pengembangan perangkat lunak *data mining* dan memberikan wawasan tentang masa depan *data mining*, termasuk tren dan perkembangan terbaru dalam teknologi ini. Dengan studi kasus yang mendalam, pembaca dapat melihat aplikasi praktis dari perangkat lunak *data mining* dalam berbagai sektor industri.

Buku ini dirancang untuk menjadi sumber daya yang bermanfaat bagi yang ingin mempelajari *Data Mining*, dari analisis data hingga pengambil keputusan strategis, serta menambah pengetahuan praktis tentang bagaimana memaksimalkan potensi perangkat lunak *Data Mining* dalam berbagai sektor industri.

