

Web Mining.docx

by Student Turnitin

Submission date: 27-Aug-2025 11:04AM (UTC-0700)

Submission ID: 2736352810

File name: Web_Mining.docx (17.87K)

Word count: 419

Character count: 2774

1 Web Mining

Web Mining adalah penerapan teknik data mining untuk menggali informasi, pola, atau pengetahuan dari data yang ada di web. Data web bisa berupa konten halaman (teks, gambar, video), struktur link antar halaman, maupun jejak penggunaan pengguna (usage data). Tujuannya adalah memanfaatkan data web untuk berbagai kepentingan, seperti bisnis, sistem rekomendasi, analisis perilaku pengguna, maupun pengembangan teknologi web yang lebih pintar.

Komponen Penting dalam Web Mining

1. Web Scraping

Teknik untuk mengambil data dari halaman web secara otomatis menggunakan program/script. **Tujuan:** mengumpulkan informasi yang ditampilkan di web (teks, tabel, harga produk, artikel). **Contoh:** mengambil data harga barang dari e-commerce untuk perbandingan harga.

2. Web Crawling

Proses menelusuri dan mengindeks banyak halaman web dengan bantuan program yang disebut crawler/spider. **Tujuan:** mengumpulkan data web dalam skala besar untuk dianalisis/disimpan. **Contoh:** Googlebot yang menjelajahi web untuk memperbarui indeks mesin pencari Google.

3. Web Data Preprocessing

Tahap persiapan data web sebelum dilakukan analisis. **Langkah-langkah:**

- Data Cleaning → menghapus data yang tidak relevan/rusak.
- Deduplication → menghilangkan data ganda.
- Normalization → menyeragamkan format data.
- Transformation → mengubah data agar lebih mudah diproses.

Tujuan: membuat data lebih berkualitas sehingga hasil analisis lebih akurat.

4. Pembelajaran Terawasi (Supervised Learning)

Teknik machine learning dengan data berlabel (sudah ada target/jawaban). **Cara kerja:** model belajar dari data latih, lalu memprediksi data baru. **Contoh algoritma:** Decision Tree, SVM, Neural Network. **Penerapan:** klasifikasi email spam, prediksi harga rumah.

5. Pembelajaran Tak Terawasi (Unsupervised Learning)

Teknik machine learning dengan data tanpa label. **Tujuan:** menemukan pola atau pengelompokan dalam data. **Contoh algoritma:** K-Means, Association Rules, PCA. **Penerapan:** mengelompokkan pelanggan berdasarkan kebiasaan belanja, analisis log pengguna.

Taksonomi Web Mining

1. Web Structure Mining

- Mining XML (DTD Schema) → analisis struktur data berbasis XML.
- Mining HTML Documents → menggali informasi dari struktur HTML.
- Navigation → menganalisis pola navigasi pengguna melalui hyperlink.

2. Web Usage Mining

- Personalization → menyesuaikan layanan sesuai preferensi pengguna.
- Business Intelligence → mendukung pengambilan keputusan bisnis.
- Customer/User Profile → membentuk profil pengguna dari perilaku.
- System Improvement → meningkatkan performa sistem web.
- Recommendation → sistem rekomendasi (contoh: e-commerce).
- E-Commerce → penerapan mining pada transaksi online.
- Intrusion Detection → mendeteksi akses mencurigakan/serangan.
- Web Agents → agen otomatis untuk analisis aktivitas web.

3. Web Content Mining

- Clustering → mengelompokkan dokumen web yang mirip.
- Association Rule → menemukan hubungan antar data.
- Semantic Web → memahami makna data web agar lebih mudah diproses.
- Web Page Content Mining → menggali isi halaman (teks, link, metadata).
- Search Result Mining → menganalisis hasil pencarian.
- Text Mining → menggali informasi dari data teks.
- Image Mining → menggali informasi dari data gambar.

Web Mining.docx

ORIGINALITY REPORT

3%

SIMILARITY INDEX

3%

INTERNET SOURCES

0%

PUBLICATIONS

0%

STUDENT PAPERS

PRIMARY SOURCES

1

nandayolanda.blogspot.co.id

Internet Source

3%

Exclude quotes Off

Exclude bibliography Off

Exclude matches Off