

Analisis Data Alergi dan Katering Sekolah Oktober 2024
Menggunakan Data Mining



Disusun oleh :

Robithoh Alam – A11.2022.14210

JURUSAN TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS DIAN NUSWANTORO

Deskripsi Singkat

Makalah ini membahas penerapan data mining dalam analisis data alergi dan absensi siswa terkait layanan katering di sebuah sekolah. Tujuannya adalah untuk menemukan pola atau insight yang berguna, seperti kaitan antara alergi makanan dan kehadiran, serta keterlambatan pembayaran katering.

Masalah dan Tujuan yang akan di capai

1. Masalah

- a. Identifikasi pengaruh alergi makanan terhadap absensi siswa dalam layanan katering.
- b. Pola keterlambatan pembayaran dalam layanan katering dan faktor-faktor yang mempengaruhinya.
- c. Korelasi antara status siswa (misalnya, anak guru) dengan perilaku pembayaran.

2. Tujuan

- a. Menganalisis hubungan antara jenis alergi dan tingkat kehadiran.
- b. Mendeteksi pola pada siswa yang belum melunasi pembayaran dan faktor yang mungkin memengaruhi.
- c. Mengidentifikasi pola perilaku siswa terkait absensi dan status pembayaran berdasarkan status sosial-ekonomi (misalnya, anak guru).

Penjelasan Dataset

Dataset yang digunakan dalam eksperimen ini terdiri dari informasi mengenai siswa dan layanan katering, termasuk:

1. Kode: Kode unik setiap siswa.
2. Kelas: Kategori kelas yang diikuti oleh siswa.
3. Nama: Nama siswa.
4. Jenis Alergi: Terdiri dari berbagai jenis alergi seperti alergi telur puyuh, ikan pindang, daging sapi, dan lainnya.
5. Absensi (Total Tidak Masuk/Total Masuk): Menunjukkan jumlah hari siswa tidak hadir dan hadir dalam layanan katering.
6. Keterangan: Informasi tambahan seperti siswa anak guru.
7. Total Pembayaran: Jumlah yang harus dibayar oleh siswa untuk layanan catering.
8. Status Pembayaran: Status pembayaran apakah lunas, belum lunas, atau ada keterangan khusus.

Fitur Utama dalam Dataset

1. Jenis Alergi: Berisi boolean untuk setiap jenis alergi.
2. Absensi Katering: Total jumlah hari siswa tidak mengikuti layanan katering.
3. Status Pembayaran: Status pembayaran, baik lunas, belum lunas, atau keterangan lain.
4. Siswa Anak Guru: Indikator apakah siswa merupakan anak guru yang mendapatkan diskon.

Alur

1. Pengumpulan Data: Data yang dikumpulkan dari layanan katering sekolah terkait absensi siswa, jenis alergi, dan status pembayaran.
2. Preprocessing Data:
 - a. Cleaning data dari nilai kosong atau duplikat.
 - b. Transformasi fitur alergi menjadi kategori biner (TRUE/FALSE).
 - c. Normalisasi data pembayaran dan absensi
3. Eksplorasi data
 - a. Melakukan analisis deskriptif seperti rata-rata absensi per kelas, dan frekuensi siswa dengan alergi tertentu.
 - b. Analisis distribusi pembayaran dan statusnya.
4. Pemodelan data
 - a. Menggunakan metode klasifikasi seperti Decision Tree atau Random Forest untuk memprediksi status pembayaran berdasarkan fitur seperti absensi, alergi, dan status siswa.
 - b. Menggunakan analisis korelasi untuk memahami hubungan antara alergi dan tingkat absensi.
5. Evaluasi
 - a. Mengukur akurasi prediksi status pembayaran.
 - b. Mengidentifikasi insight dari hasil analisis, misalnya, siswa dengan alergi tertentu lebih sering absen atau keterlambatan pembayaran berkorelasi dengan absensi yang tinggi.
6. Interpretasi hasil dan visualisasi
 - a. Menyajikan hasil dalam bentuk tabel, grafik, dan diagram untuk memudahkan pemahaman.
 - b. Menyusun laporan akhir yang mencakup rekomendasi berdasarkan hasil analisis.

Timeline Eksperimen

Tanggal	Tahap Eksperimen	Deskripsi kegiatan
18 Oktober 2024	Pengenalan Dataset & Pemahaman Data	Mengumpulkan dan memahami dataset siswa yang diberikan, termasuk fitur-fitur seperti alergi, absensi, dll.
20 Oktober 2024	Pembersihan Data	Mengidentifikasi dan menangani nilai yang hilang, data duplikat, atau data yang tidak valid.
22 Oktober 2024	Preprocessing Data	Melakukan transformasi data: normalisasi, encoding untuk fitur kategorikal (Status Pembayaran, Jenis Alergi)
24 Oktober 2024	Exploratory Data Analysis (EDA)	Melakukan analisis eksplorasi data menggunakan visualisasi untuk melihat pola umum dan distribusi data.
27 Oktober 2024	Feature Engineering	Menciptakan fitur-fitur baru berdasarkan kombinasi variabel yang ada (misalnya, total ketidakhadiran).
29 Oktober 2024	Pemilihan Model & Algoritma	Memilih algoritma yang cocok untuk masalah yang akan diselesaikan (misalnya, klasifikasi status pembayaran).
31 Oktober 2024	Pembagian Data & Validasi	Membagi dataset menjadi data latih dan data uji, serta menentukan metode validasi (cross-validation).
2 November 2024	Pelatihan Model	Melatih model menggunakan algoritma yang dipilih pada data latih.
4 November 2024	Evaluasi Model	Mengevaluasi performa model menggunakan metrik seperti akurasi, precision, recall, dll.
5 November 2024	Optimasi Model	Melakukan tuning parameter untuk meningkatkan performa model.

6 November 2024	Interpretasi Hasil dan Penyimpulan	Menganalisis hasil akhir, memberikan kesimpulan, serta memberikan rekomendasi berdasarkan analisis.
-----------------	---------------------------------------	---