

**LAPORAN TUGAS
JARINGAN SYARAF TIRUAN**

**IMPLEMENTASI SELF ORGANIZING-MAPS (SOM)
PADA SEGMENTASI PELANGGAN**



Disusun Oleh:

Qolbu Salim NIM. 23031030020

Dosen Pengampu:

Nur Insani, Ph.D.

**PROGRAM STUDI STATISTIKA
DEPARTEMEN PENDIDIKAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI YOGYAKARTA
2025**

DESKRIPSI DATA

Sumber

Data yang digunakan dalam analisis ini adalah data sekunder yang dengan judul *Online Retail* yang berasal dari UCI Machine Learning Repository (ID: 352). Dataset ini berisi transaksi retail online yang terjadi antara Desember 2010 dan Desember 2011 untuk perusahaan retail non-toko yang berbasis di Inggris. Perusahaan ini terutama menjual hadiah unik untuk berbagai acara, dan banyak pelanggannya adalah grosir.

Karakteristik

Data ini berisi total 541.909 baris dengan 8 kolom, dimana kolom tersebut terdiri dari 6 kolom fitur dan 2 kolom id. Data ini memiliki karakteristik multivariat, runtun waktu, dan sekuensial. Dataset ini tidak memiliki variabel target, yang berarti data ini cocok untuk kasus *unsupervised learning*.

Variabel

| Nama Variabe | Keterangan |
|--------------|--|
| InvoiceNo | Nomor invoice (kategori) |
| StockCode | Kode produk (kategori) |
| Description | Deskripsi produk (kategori) |
| Quantity | Kuantitas produk per transaksi (numerik) |
| InvoiceDate | Tanggal dan waktu transaksi (date) |
| UnitPrice | Harga per unit dalam sterling (numerik) |
| CustomerID | ID pelanggan (kategori) |
| Country | Negara asal pelanggan (kategori) |

DESKRIPSI ALGORITMA

Penjelasan Parameter:

1. Map Size

Ukuran peta (*map size*) pada *Self-Organizing Map* (SOM) merupakan parameter yang menentukan jumlah dan susunan neuron dalam grid dua dimensi yang digunakan untuk merepresentasikan data. Pemilihan ukuran peta dilakukan melalui proses evaluasi kualitas kluster dengan mempertimbangkan dua metrik utama, yaitu *Quantization Error* (QE) dan *Topographic Error* (TE). *Quantization Error* mengukur seberapa dekat data dengan neuron terbaiknya (*Best Matching Unit*), sedangkan *Topographic Error* mengindikasikan sejauh mana struktur topologi data dapat dipertahankan dalam representasi peta. Dalam penelitian ini, ukuran peta yang diuji bervariasi antara 5×5

hingga 15×15 . Ukuran optimal ditentukan berdasarkan keseimbangan antara nilai QE yang rendah yang menandakan representasi data yang akurat dan nilai TE yang kecil yang menunjukkan preservasi struktur topologi yang baik.

2. *Sigma*

Parameter *sigma* berfungsi untuk menentukan radius dari fungsi *neighborhood* pada saat proses pelatihan SOM. Nilai sigma mengatur seberapa luas pengaruh neuron pemenang terhadap neuron-neuron di sekitarnya. Nilai sigma yang kecil akan menghasilkan *neighborhood* yang sempit, sehingga pembaruan bobot hanya berdampak pada neuron yang sangat berdekatan dengan neuron pemenang. Sebaliknya, nilai sigma yang besar menciptakan *neighborhood* yang luas, sehingga lebih banyak neuron yang turut diperbarui dalam setiap iterasi. Pada penelitian ini, nilai sigma ditetapkan sebesar 1,5, yang dianggap cukup untuk menjaga keseimbangan antara kemampuan generalisasi dan ketepatan representasi lokal.

3. *Learning Rate*

Parameter *learning rate* menunjukkan tingkat pembelajaran yang mengatur besarnya langkah pembaruan bobot (*weight update*) selama proses pelatihan SOM. Nilai ini mengontrol seberapa besar perubahan yang diterapkan pada bobot neuron setiap kali data baru diproses. Nilai *learning rate* yang terlalu tinggi dapat menyebabkan pelatihan menjadi tidak stabil, sedangkan nilai yang terlalu rendah dapat memperlambat proses konvergensi. Pada implementasi ini, nilai *learning rate* ditetapkan sebesar 0,5, yang memberikan keseimbangan antara stabilitas pembelajaran dan kecepatan konvergensi menuju kondisi peta yang terorganisasi dengan baik.

4. *Number of Iterations*

Jumlah iterasi menunjukkan berapa kali proses pelatihan dilakukan untuk menyesuaikan bobot jaringan terhadap distribusi data. Setiap iterasi mencakup pemilihan sampel data, identifikasi neuron pemenang, dan pembaruan bobot berdasarkan fungsi *neighborhood* serta *learning rate* yang berlaku. Dalam konteks penelitian ini, jumlah iterasi ditetapkan sebanyak 10.000, yang dianggap cukup untuk mencapai konvergensi model tanpa menyebabkan *overfitting*. Jumlah iterasi tersebut memastikan bahwa SOM memiliki waktu yang memadai untuk menyesuaikan diri terhadap pola data dan menghasilkan peta yang stabil serta bermakna.

5. *Random Seed*

Parameter *random seed* digunakan untuk menjamin *reproducibility* atau keterulangan hasil pelatihan model. Karena proses inisialisasi bobot awal pada SOM bersifat acak, penetapan *random seed* memungkinkan hasil eksperimen dapat direplikasi dengan kondisi yang sama pada setiap pelatihan ulang. Dengan menetapkan *random seed* sebesar 42, proses pelatihan SOM dapat menghasilkan representasi peta yang konsisten dan dapat dibandingkan secara objektif antar percobaan.

HASIL DAN PEMBAHASAN

Penerapan *Self-Organizing Map* (SOM) pada dataset *Online Retail* berhasil mengidentifikasi beberapa kluster utama pelanggan dengan profil perilaku berbeda, yang divisualisasikan melalui analisis U-Matrix dan Customer Distribution Heatmap. Konsentrasi

signifikan pelanggan ditemukan di sisi kanan peta, menunjukkan kelompok homogen dengan aktivitas tinggi dan nilai ekonomi besar, sementara sisi kiri dan bawah didominasi oleh pelanggan dengan aktivitas lebih rendah dan sensitivitas harga yang tinggi. Visualisasi fitur memperlihatkan pola spasial konsisten yang mencerminkan keragaman perilaku pelanggan di seluruh grid.

Penelitian ini membuktikan efektivitas SOM dalam memetakan pola perilaku pelanggan secara menyeluruh, memudahkan identifikasi segmen bernilai tinggi, menengah, dan rendah. Temuan tersebut menggarisbawahi pentingnya diferensiasi strategi pemasaran berdasarkan karakteristik segmen, pengembangan program loyalitas yang fokus, serta komunikasi yang efisien untuk segmen bernilai ekonomis terbatas. Hasil ini menyediakan dasar empiris kuat untuk pengambilan keputusan strategis dalam mengoptimalkan alokasi sumber daya pemasaran dan meningkatkan nilai pelanggan secara berkelanjutan.

SUMBER KODE

Kode dan dokumentasi pada analisis ini disimpan ke dalam *Repository* Github dengan tautan berikut:

<https://github.com/qlbusalim/Self-Organizing-Map-SOM>

REFERENSI

<https://github.com/JustGlowing/minisom>

<https://github.com/diego-vicente/som-tsp>

Kohonen, T. (2001). Self-Organizing Maps

Alie, J., & Gustriansyah, R. (2024). Customer Segmentation For Digital Marketing Based on Shopping Patterns. *Jurnal Aplikasi Bisnis Dan Manajemen*, 10(1), 209. <https://doi.org/10.17358/jabm.10.1.209>

Sayan, İsmail & Demirdağ, Melike & Yüçetürk, Güven & Yalçınkaya, Sare. (2022). A Review of Customer Segmentation Methods: The Case of Investment Sector. 200-204. 10.1109/BDAI56143.2022.9862801.

Asan, Umut & Ercan, Secil. (2012). An Introduction to Self-Organizing Maps. 10.2991/978-94-91216-77-0_14.