TUGAS PERTEMUAN KE 3

REGERESI LOGISTIK

Tugas:

1. Buat Model untuk Regresi Logistik
2. Metrik Pengukuran Regresi Logistik + Confusion Matrix (Uraikan)
3. Kelebihan dan Kekurangan Regresi Logistik
4. Cari 10 Jurnal terkait pemanfaatan Regresi Logistik
5. Diskusikan dalam Forum
6. Tuliskan dalam laporan (dikumpulkan saat UTS)
7. **Buat Model untuk Regresi Logistik**
   1. Deskripsi Data

Seorang peneliti ingin mengetahui bagaimana pengaruh kualitas pelayanan publik

terhadap kepuasan pengguna (masyarakat). Kualitas pelayanan publik diteliti

melaluji variabel Daya Tanggap (X1) dan Empati (X2). Kepuasan penggunana

layanan (Y) sebagai variabel dependent adalah variabel dummy dimana dimana

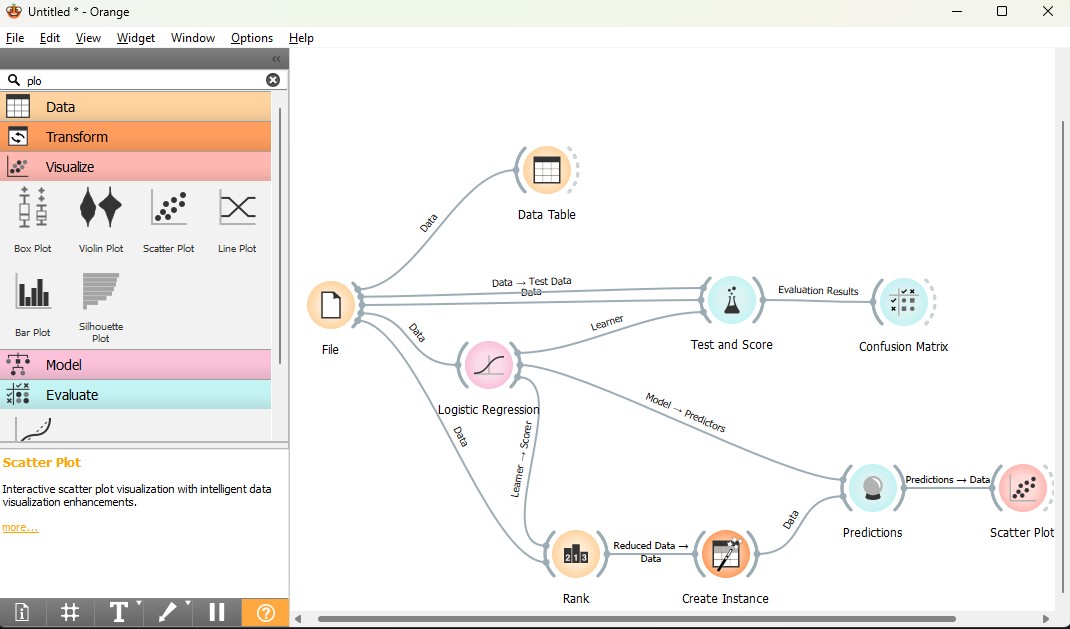
jika responden menjawab puas maka kita beri skor 1 dan jika menjawab tidak puas

kita beri skor 0.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Person** | **Daya Tanggap (X1)** | **Empati (X2)** | **Kepuasan Pelanggan (Y)** |  | **Person** | **Daya Tanggap (X1)** | **Empati (X2)** | **Kepuasan Pelanggan (Y)** |
| 1 | 36 | 45 | 1 |  | 26 | 33 | 41 | 0 |
| 2 | 34 | 39 | 0 |  | 27 | 32 | 39 | 0 |
| 3 | 30 | 38 | 0 |  | 28 | 30 | 36 | 0 |
| 4 | 32 | 38 | 1 |  | 29 | 30 | 36 | 0 |
| 5 | 36 | 45 | 1 |  | 30 | 36 | 42 | 1 |
| 6 | 33 | 42 | 0 |  | 31 | 33 | 38 | 0 |
| 7 | 36 | 45 | 1 |  | 32 | 33 | 38 | 0 |
| 8 | 36 | 45 | 1 |  | 33 | 35 | 41 | 1 |
| 9 | 31 | 36 | 0 |  | 34 | 35 | 41 | 1 |
| 10 | 31 | 37 | 0 |  | 35 | 34 | 40 | 1 |
| 11 | 36 | 45 | 1 |  | 36 | 30 | 38 | 0 |
| 12 | 33 | 41 | 0 |  | 37 | 30 | 40 | 1 |
| 13 | 32 | 40 | 0 |  | 38 | 35 | 41 | 1 |
| 14 | 33 | 39 | 0 |  | 39 | 34 | 42 | 1 |
| 15 | 34 | 42 | 1 |  | 40 | 33 | 40 | 0 |
| 16 | 34 | 42 | 0 |  | 41 | 34 | 43 | 0 |
| 17 | 32 | 39 | 0 |  | 42 | 30 | 38 | 0 |
| 18 | 34 | 42 | 1 |  | 43 | 34 | 42 | 1 |
| 19 | 33 | 40 | 0 |  | 44 | 30 | 41 | 1 |
| 20 | 34 | 43 | 0 |  | 45 | 34 | 40 | 0 |
| 21 | 32 | 39 | 0 |  | 46 | 34 | 42 | 1 |
| 22 | 36 | 44 | 1 |  | 47 | 34 | 38 | 0 |
| 23 | 33 | 37 | 0 |  | 48 | 34 | 44 | 1 |
| 24 | 30 | 38 | 0 |  | 49 | 35 | 43 | 0 |
| 25 | 36 | 43 | 1 |  | 50 | 34 | 42 | 1 |

* 1. Pengolahan Data

Pengolahan data menggunakan Orange

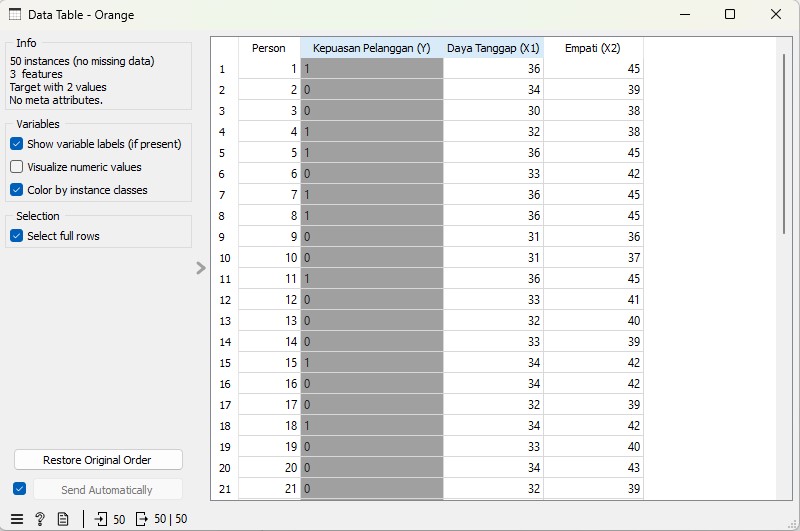


Gambar 1.2 Gambar Pengolahan Data

* 1. Intepretasi Data

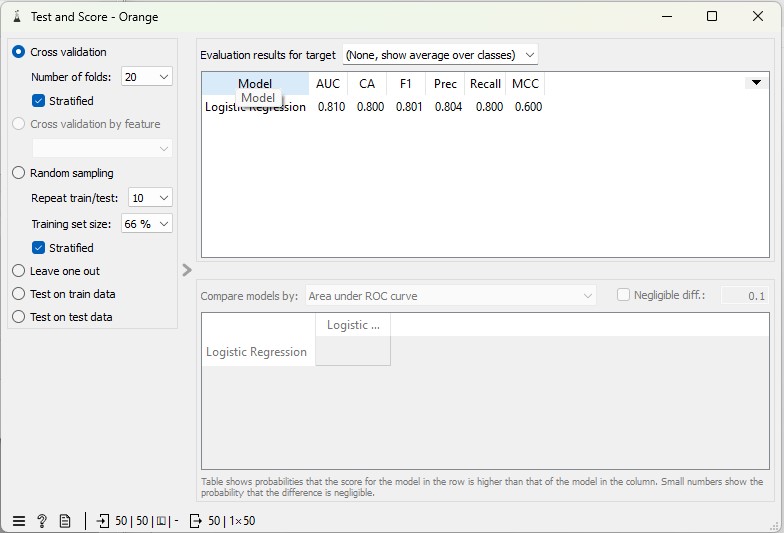
Berikut penjelsan output dari pengolaha data

2



Gambar 1.3 Output data

Dari gambar diatas jumlah data sebanyak 50 dan target yang akan dijadikan target adalah feature kepuasaan dengan data numerical, dan daya tanggap juga empati sebagai x1 dan x2.

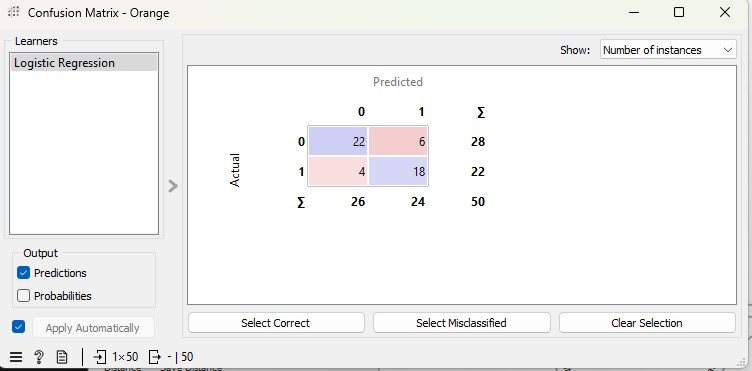


Gambar 1.3 Output nilai AUC , CA,f1 Prec Recall

Dari output diatas didapat nilai:

* AUC = 0.810 ini artinya ,
* CA = 0.80 ini artinya,
* Prec = 0.804 ini berarti 80% model akurat dalam mengklasifikasikan dengan benar dengan yang diprediksi
* Recall 0.8 artinya 80% menggambarkan keberhasilan model dalam menemukan kembali sebuah informasi
* MCC 0.6

1. **Metrik Pengukuran Regresi Logistik + Confusion Matrix (Uraikan)**



Gambar 1.4 Confusion Matrix

Dari Confusion Matrix dilihat ada nilai actual positif/true positif (TP) bernilai 22 orang secara actual merasa puas d, tetapi nilai prediksi negative (TN) benilai 6 artinya diprediksi ada 6 orang merasa tidak puas a, dan 4 orang secara actual merasa tidak puas tetapi diprediksi merasa puas (FN) dan nilai prediksi 18 orang tidak merasa puas dan hal ini sesuai dengan actual (FP)

1. **Kelebihan dan Kekurangan Regresi Logistik**

Kelebihan:

* Cocok untuk data biner: Regresi logistik sangat berguna saat Anda bekerja dengan variabel dependen yang bersifat biner atau kategorikal.
* Interpretasi yang Mudah: Hasil dari regresi logistik (koefisien) dapat diinterpretasikan secara langsung dalam hal log-odds atau peluang (odds ratio), membuatnya lebih mudah dipahami.
* Kemampuan untuk Menangani Variabel Independen: Regresi logistik dapat menangani berbagai jenis variabel independen, termasuk kategorikal dan kontinu.
* Mendeteksi Hubungan Non-linear: Meskipun regresi logistik adalah model linier, Anda bisa menambahkan fitur non-linear dengan menambahkan polinomial atau menggunakan transformasi variabel.
* Mengatasi Multikolinearitas: Regresi logistik lebih tahan terhadap masalah multikolinearitas (ketergantungan antar variabel independen) dibandingkan dengan regresi linier.

Kekurangan:

* Asumsi tentang Independensi: Regresi logistik mensyaratkan asumsi independensi antar pengamatan, artinya pengamatan tidak boleh saling bergantung. Ini bisa menjadi masalah jika data yang Anda miliki memiliki ketergantungan temporal atau spasial.
* Rentan terhadap Overfitting: Seperti model statistik lainnya, regresi logistik dapat menjadi rentan terhadap overfitting jika Anda tidak berhati-hati dalam pemilihan fitur atau penyesuaian model.
* Tidak Cocok untuk Variabel Dependensi Kontinu: Regresi logistik dirancang khusus untuk variabel dependen biner, sehingga tidak dapat digunakan langsung untuk variabel dependen yang kontinu.
* Keterbatasan dalam Kompleksitas Model: Regresi logistik merupakan model linier, sehingga tidak dapat menangani hubungan yang sangat kompleks antara variabel independen dan dependen. Ini bisa menjadi kendala jika hubungan tersebut bersifat non-linier.
* Dibutuhkan Data yang Besar: Regresi logistik sering memerlukan jumlah pengamatan yang cukup besar untuk menghasilkan hasil yang dapat diandalkan. Jika jumlah sampel terlalu kecil, estimasi parameter dapat menjadi tidak stabil