



PART 9

Modeling Algorithm and Model Evaluation

VERSI 2025

Modeling Algorithm

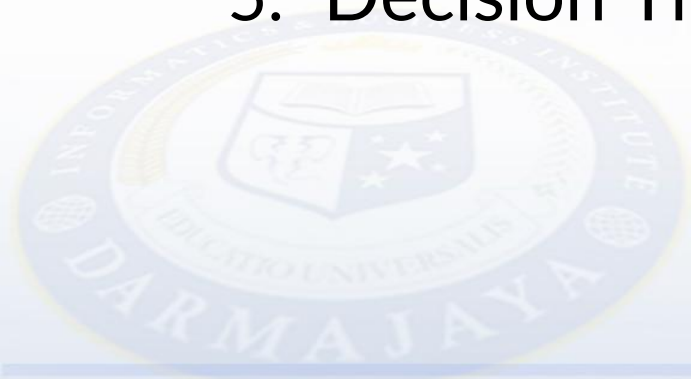
Business Intelligence (BI) tidak hanya menganalisis apa yang terjadi di masa lalu, tetapi juga memprediksi *apa yang mungkin terjadi*, *siapa target yang tepat*, dan *bagaimana strategi terbaik*.

Algoritma klasifikasi sangat penting untuk mendukung keputusan tersebut.



Algoritma Klasifikasi:

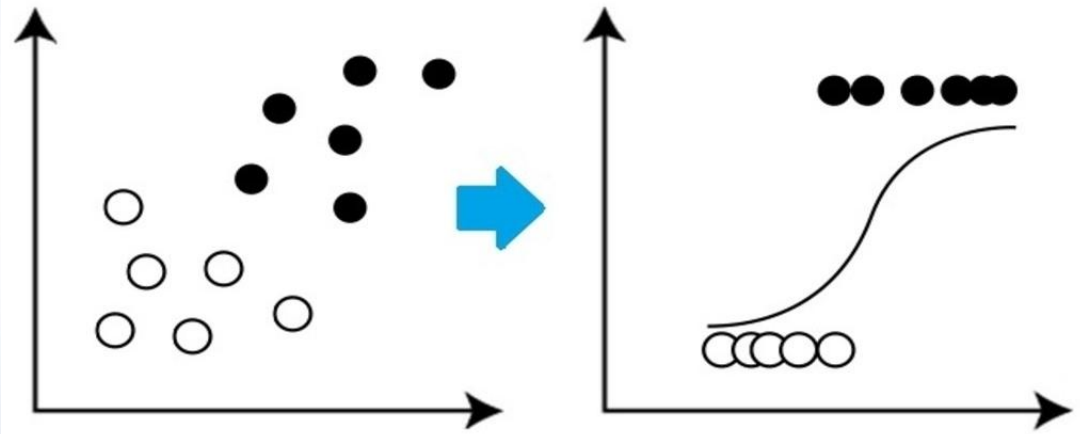
1. Logistic regression
2. K-Nearest Neighbors
3. Naive Bayes
4. Support Vector Machines
5. Decision Tree



Algoritma Logistic regression:

Logistic Regression adalah algoritma klasifikasi yang sangat mendasar namun penting dalam machine learning yang menggunakan satu atau lebih variabel independen untuk menentukan suatu hasil. Logistic Regression mencoba menemukan hubungan yang paling sesuai antara variabel terikat dan sekumpulan variabel bebas.

- Logistic Regression membantu perusahaan memprediksi dua kemungkinan keputusan (yes/no, beli/tidak beli, respon/tidak respon).



Contoh Penggunaan Bisnis

1. **Memprediksi apakah pelanggan akan membeli produk promo**

Input: umur, penghasilan, riwayat belanja

Hasil: probabilitas pembelian promo

2. **Memprediksi risiko churn (pelanggan berhenti membeli)**

Perusahaan bisa melakukan retargeting lebih cepat.

3. **Mengukur faktor yang paling memengaruhi keputusan belanja**

LR memberikan “koefisien” sehingga perusahaan tahu faktor paling ku

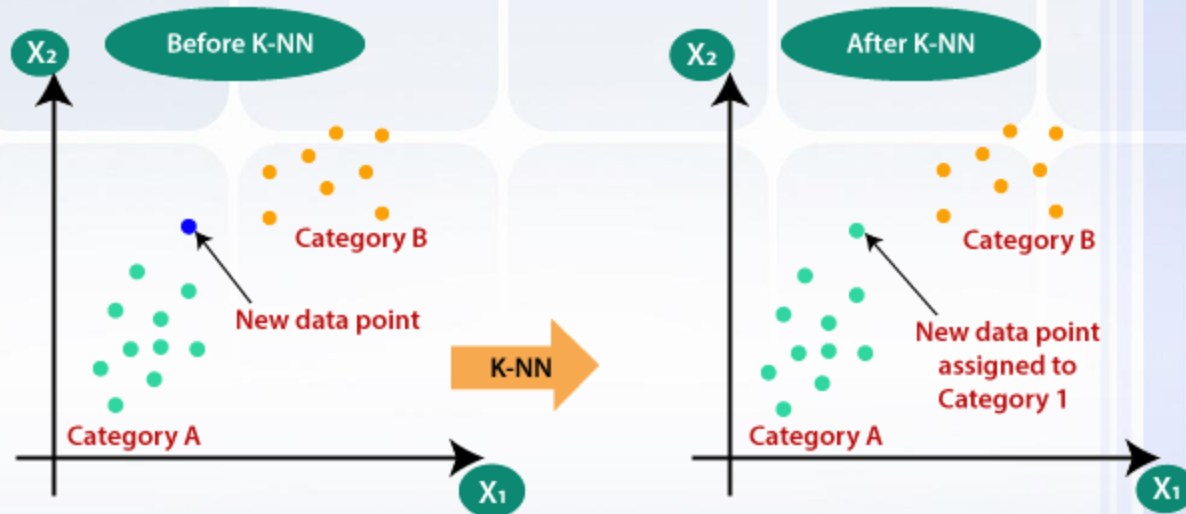
Dampak Bisnis

- Mempersempit target promosi → biaya marketing turun
- Menghindari kampanye yang tidak efektif
- Peningkatan konversi 15–30% pada kampanye targeted

Algoritma K-Nearest Neighbors

KNN mengklasifikasikan titik data baru tergantung pada kelas dari sebagian besar titik data di antara K Neighbour, dimana K adalah jumlah Neighbour yang akan dipertimbangkan. KNN menangkap gagasan kesamaan dengan beberapa rumus jarak matematika dasar.

KNN digunakan untuk memahami kesamaan perilaku antar pelanggan.



Contoh Penggunaan Bisnis

- **Rekomendasi Produk “Pelanggan Mirip Anda Membeli...”**
Jika pelanggan A mirip dengan pelanggan B, maka produk yang dibeli B dapat direkomendasikan ke A.
- **Segmentasi Pelanggan Berdasarkan Perilaku**
KNN mengelompokkan pelanggan berdasarkan:
 - frekuensi belanja
 - jenis produk
 - total transaksi
- **Deteksi anomali perilaku transaksi**
Mendukung fraud detection.

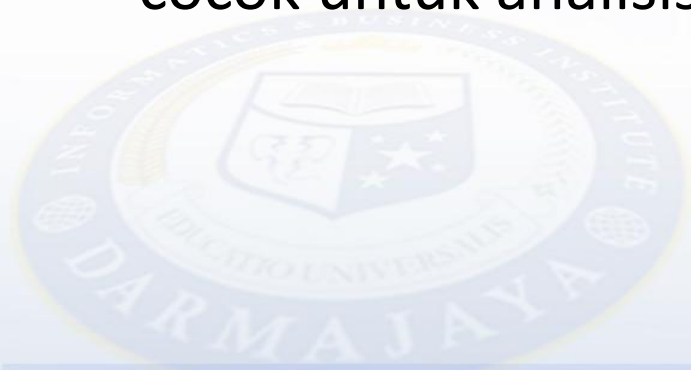
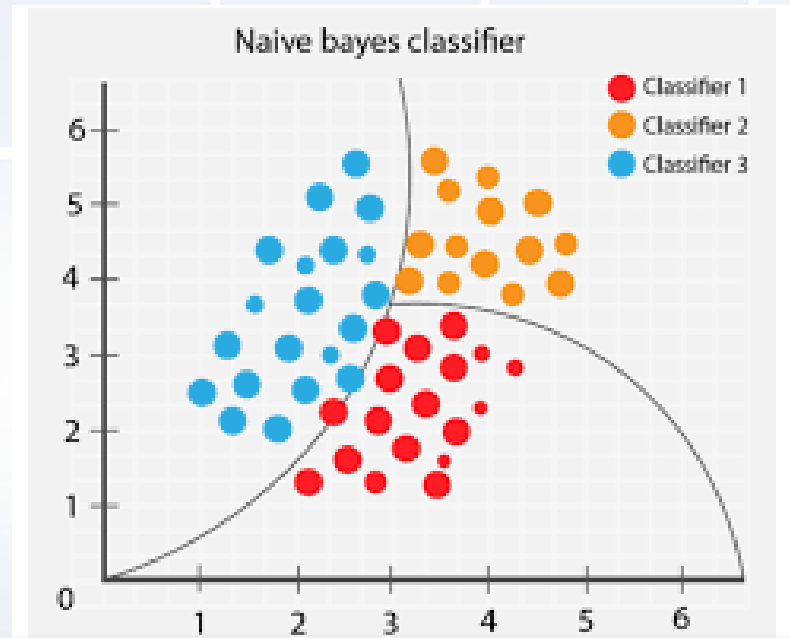
Dampak Bisnis

- Rekomendasi lebih akurat
- Cross-selling meningkat (12–28%)
- Retensi pelanggan naik karena rekomendasi personal

Algoritma Naive Bayes

Naive Bayes memberikan asumsi independensi antar predictor yang mengasumsikan bahwa keberadaan suatu fitur tertentu dalam suatu kelas tidak berhubungan dengan keberadaan variabel lainnya.

Naive Bayes adalah algoritma cepat dan efisien, sangat cocok untuk analisis teks.



Contoh Penggunaan Bisnis

1. Analisis sentimen ulasan pelanggan

Mengidentifikasi apakah review positif/negatif.

2. Klasifikasi email atau chat pelanggan

Mengelompokkan pesan masuk berdasarkan kategori:

1. komplain
2. permintaan harga
3. permintaan refund

3. Customer satisfaction prediction

Dari komentar CS + review marketplace.

Dampak Bisnis

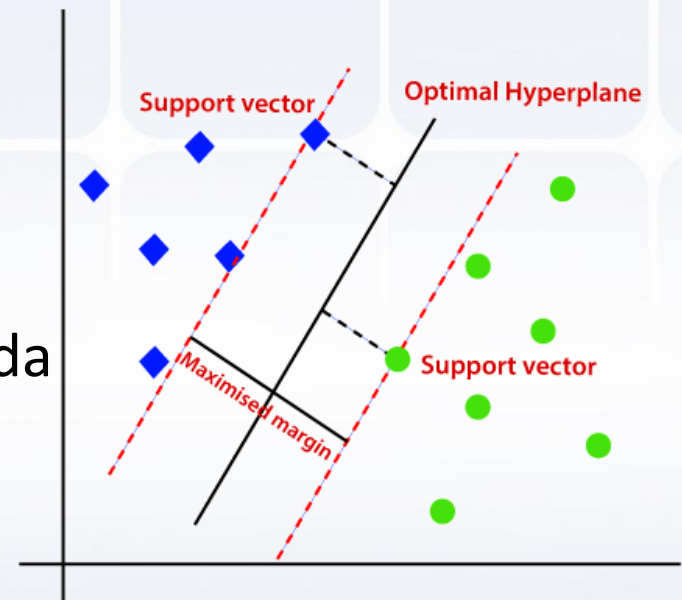
- Respon lebih cepat
- Deteksi masalah produk lebih awal
- Strategi promosi berdasarkan sentimen publik

Algoritma Support Vector Machines

SVM bekerja untuk mencari hyperplane atau fungsi pemisah (decision boundary) terbaik untuk memisahkan dua buah kelas atau lebih pada ruang input. Hiperplane dapat berupa line atau garis pada dua dimensi dan dapat berupa flat plane pada multiple plane.

SVM kuat dalam memisahkan data kompleks. Berguna untuk:

- prediksi perilaku pelanggan
- segmentasi kelas yang sangat berbeda
- deteksi fraud



Contoh Penggunaan Bisnis

1.Memprediksi pelanggan mana yang paling berpotensi membeli produk mahal

SVM memisahkan “big spender vs normal spender”.

2.Mendeteksi transaksi yang kemungkinan fraud

Berbasis pola transaksi mencurigakan.

3.Segmentasi pelanggan high-value

Membantu menentukan prioritas promosi.

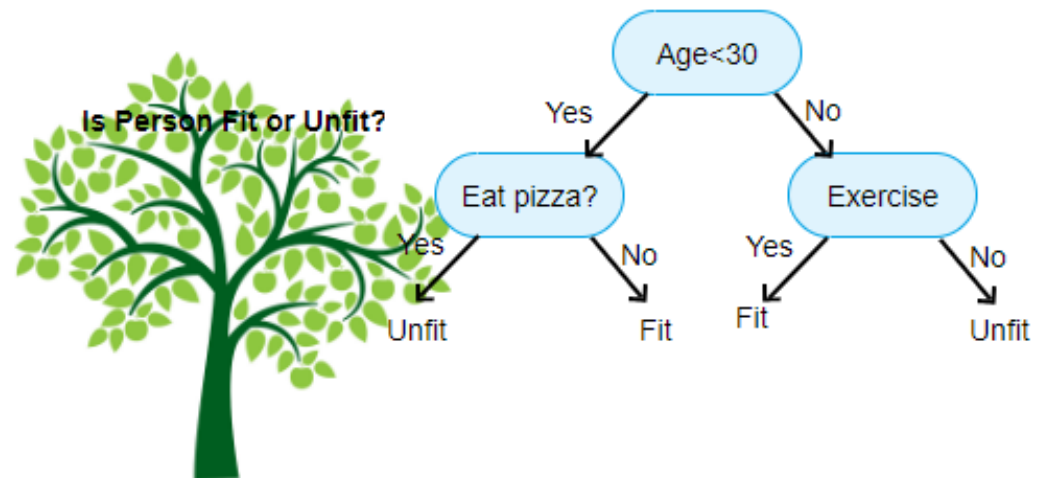
Dampak Bisnis

- ROI kampanye promosi meningkat
- Pengurangan kerugian akibat fraud
- Penargetan iklan lebih presisi

Algoritma Decision Tree

Decision Tree digunakan untuk masalah klasifikasi dan regresi dengan cara meniru pemikiran tingkat manusia sehingga sangat mudah untuk memahami data dan membuat intuisi dan interpretasi yang baik. Decision Tree benar-benar membuat pengguna melihat logika untuk menafsirkan data.

Decision Tree sangat mudah dipahami dan cocok untuk membuat keputusan bisnis berbasis aturan.



Contoh Penggunaan Bisnis

1. Menentukan strategi promosi berdasarkan perilaku pelanggan

Contoh rule:

1. Jika pembelian $< 3x$ dan umur $< 25 \rightarrow$ promo TikTok
2. Jika pembelian $> 10x \rightarrow$ loyalty voucher

2. Analisis faktor yang paling memengaruhi pembelian

DT menunjukkan:

1. harga
2. diskon
3. lokasi
4. hari pembelian

3. Prediksi keberhasilan kampanye marketing

“Kapan kampanye lebih sukses?” \rightarrow berdasarkan dataset tahun sebelumnya.

Dampak Bisnis

- Perusahaan mudah membaca logika keputusan
- Strategi marketing lebih tepat
- Bisa langsung diterapkan tanpa pengetahuan ML mendalam

Model Evaluation

Mengevaluasi model analisis sentimen dapat dilakukan dalam berbagai bentuk, bergantung pada jenis dan tingkat analisis, data dan domain, serta tujuan dan harapan pengguna. Confusion Matrix menjadi acuan Dalam mengevaluasi performance algoritma dari Machine Learning.



Predicted Class

$$\text{F1-Score} = \frac{2 * (\text{Sensitivity} * \text{Precision})}{(\text{Sensitivity} + \text{Precision})}$$

		Predicted Class		
		Positive	Negative	
Actual Class	Positive	True Positive (TP)	False Negative (FN) Type II Error	Sensitivity $\frac{TP}{(TP + FN)}$
	Negative	False Positive (FP) Type I Error	True Negative (TN)	Specificity $\frac{TN}{(TN + FP)}$
		Precision $\frac{TP}{(TP + FP)}$	Negative Predictive Value $\frac{TN}{(TN + FN)}$	Accuracy $\frac{TP + TN}{(TP + TN + FP + FN)}$

Akhir bagian Part 9

