



Institut Informatika & Bisnis  
**DARMAJAYA**  
Yayasan Alfian Husin



**Kampus  
Merdeka**  
INDONESIA JAYA

**MERDEKA  
BELAJAR**

DATA SCIENCE DARMAJAYA  
“YOUR BEST FUTURE IN DATA”

PERTEMUAN KE: 7

# **Pengenalan** *Unsupervised Learning*

**KULIAH**

OLEH: NURJOKO



# Learning Objectives

- Memahami secara fundamental apa itu unsupervised learning dan bagaimana perbedaannya dengan supervised learning
- Memahami konsep pengelompokan
- Memahami bagaimana unsupervised learning dapat diterapkan untuk mendeteksi anomali atau data yang tidak lazim dalam dataset
- Memahami pengelompokan hirarkis dan bagaimana dapat mengungkapkan struktur hirarkis dalam data
- Memahami cara memilih algoritma unsupervised learning yang paling sesuai untuk masalah tertentu.



# Pendahuluan

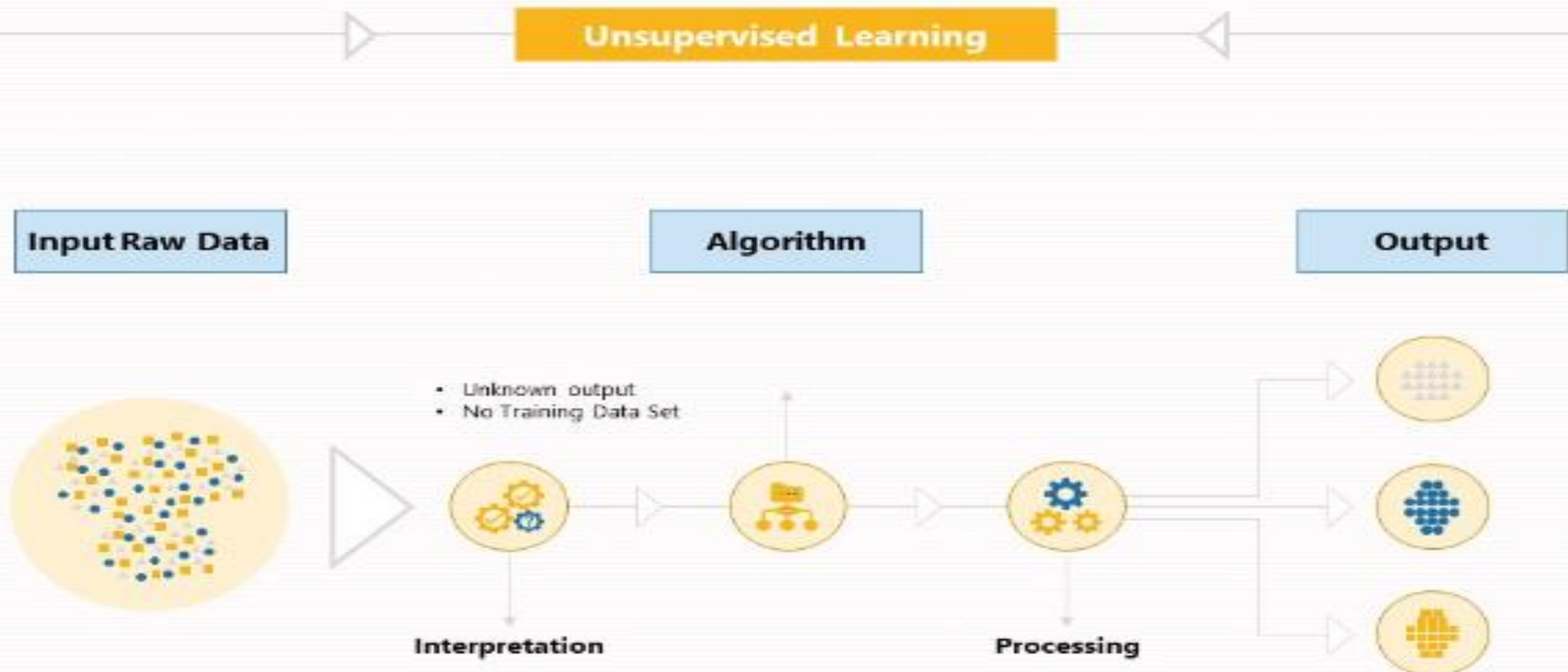
## Unsupervised Learning:

adalah salah satu paradigma dalam machine learning di mana algoritma dikonfigurasi untuk **mengekstraksi pola, struktur, dan pengetahuan dari data tanpa adanya label atau informasi target** yang telah ditentukan.

Berbeda dengan supervised learning, di mana model diberi contoh pasangan input-output untuk mempelajari hubungan, unsupervised learning menghadapi data yang tidak memiliki label



# What is Unsupervised Learning?



# Konsep Utama Dalam Unsupervised Learning

## 1. Clustering (Pengelompokan):

Salah satu tugas utama dalam unsupervised learning adalah pengelompokan data. Ini melibatkan pemisahan data menjadi kelompok (cluster) berdasarkan kesamaan fitur atau atribut. Contoh tugas clustering termasuk pengelompokan pelanggan berdasarkan perilaku pembelian atau pengelompokan dokumen berdasarkan topik.

## 2. Reduksi Dimensi:

Unsupervised learning juga digunakan untuk mereduksi dimensi data. Ini adalah proses mengurangi jumlah atribut atau fitur dalam dataset sambil mempertahankan sebanyak mungkin informasi penting. Salah satu teknik umum dalam hal ini adalah Principal Component Analysis (PCA).



### 3. Tugas Association (Asosiasi):

Tugas asosiasi dalam unsupervised learning melibatkan penemuan asosiasi atau hubungan antara item dalam dataset. Misalnya, dalam analisis belanja, tugas ini dapat digunakan untuk menemukan hubungan antara produk yang sering dibeli bersama.

### 4. Tidak Ada Label Target:

Data dalam unsupervised learning tidak memiliki label target yang harus diprediksi. Ini membedakannya dari supervised learning, di mana model dilatih untuk memprediksi output tertentu.

## 5. Overfitting:

Seperti dalam supervised learning, overfitting juga bisa menjadi masalah dalam unsupervised learning. Model dapat memisahkan data menjadi kelompok-kelompok yang sangat kecil atau tidak relevan, yang mengarah pada hasil yang tidak berguna.

## 6. Contoh Aplikasi:

Aplikasi unsupervised learning mencakup segmentasi pasar, pengelompokan dokumen, deteksi outlier, kompresi data, dan analisis faktor. Algoritma unsupervised learning membantu mengungkap pola dalam data yang mungkin tidak dapat terlihat dengan mata telanjang.

# Tipe-Tipe Masalah Unsupervised Learning

Unsupervised learning digunakan untuk berbagai tipe masalah, tergantung pada jenis tugas yang ingin dipecahkan. Berikut adalah beberapa tipe masalah utama dalam unsupervised learning:

## 1. Clustering (Pengelompokan):

Tugas clustering melibatkan pengelompokan data ke dalam kelompok berdasarkan kemiripan fitur. Kelompok-kelompok ini dapat membantu mengidentifikasi pola dalam data yang mungkin tidak terlihat sebelumnya. Contoh algoritma clustering termasuk K-Means, Hierarchical Clustering, dan DBSCAN.

## 2. Reduksi Dimensi:

Tugas reduksi dimensi bertujuan untuk mengurangi jumlah atribut atau fitur dalam dataset sambil mempertahankan informasi yang paling penting. Ini digunakan untuk mengatasi masalah dimensi tinggi dan dapat membantu menghilangkan noise dalam data. Contoh teknik reduksi dimensi termasuk Principal Component Analysis (PCA) dan t-SNE.

## 3. Anomaly Detection (Deteksi Pencurian):

Tugas ini mencakup identifikasi observasi yang tidak biasa atau langka dalam dataset. Ini berguna dalam mendeteksi aktivitas yang mencurigakan, seperti deteksi fraud dalam transaksi keuangan atau masalah dalam jaringan komputer. Algoritma untuk deteksi anomali termasuk Isolation Forest dan One-Class SVM.



#### **4. Association Rule Learning (Pembelajaran Aturan Asosiasi):**

Tugas ini melibatkan penemuan hubungan atau aturan asosiasi antara item dalam dataset. Ini digunakan dalam analisis belanja, di mana kita ingin mengetahui item apa yang sering dibeli bersama. Algoritma populer dalam tugas ini adalah Apriori.

#### **5. Latent Variable Models (Model Variabel Tersembunyi):**

Unsupervised learning juga digunakan untuk mengidentifikasi variabel tersembunyi atau laten dalam data. Ini berguna dalam pemodelan topik dalam teks atau pengelompokan data yang kompleks. Contoh termasuk Latent Dirichlet Allocation (LDA) dan autoencoders dalam deep learning.

## 7. Density Estimation (Estimasi Kepadatan):

Tugas ini mencakup estimasi distribusi probabilitas data. Ini digunakan dalam pemodelan statistik dan analisis data. Algoritma seperti Gaussian Mixture Models (GMM) digunakan dalam tugas ini.

## 8. Self-Organizing Maps (SOM):

SOM adalah jaringan saraf buatan yang digunakan untuk proyeksi data ke peta dua dimensi yang berguna dalam visualisasi data dan pengelompokan.

## 9. Pengelompokan Hirarkis (Hierarchical Clustering):

adalah tipe khusus dari tugas pengelompokan di mana kelompok-kelompok tersebut dapat membentuk hierarki yang memberikan wawasan tentang hubungan yang kompleks dalam data.



Fill in .....



Fill in .....



# CONCLUSION

Unsupervised learning memberikan alat yang kuat untuk menggali pola tersembunyi dan wawasan dari data yang belum diberi label. Ini berguna dalam berbagai aplikasi, termasuk segmentasi pelanggan, analisis teks, pengenalan pola, dan pemrosesan bahasa alami. Untuk memahami dan berhasil menerapkan unsupervised learning, penting untuk memahami berbagai algoritma dan teknik yang tersedia, serta tahu kapan dan bagaimana menggunakannya dalam konteks yang sesuai.



# REFERENCES

Fill in IEEE Style



Institut Informatika & Bisnis  
**DARMAJAYA**  
Yayasan Alfian Husin



**Kampus  
Merdeka**  
INDONESIA JAYA

**MERDEKA  
BELAJAR**

# THANK YOU!!

DATA SCIENCE DARMAJAYA "YOUR BEST FUTURE IN DATA"