



UNIT-III

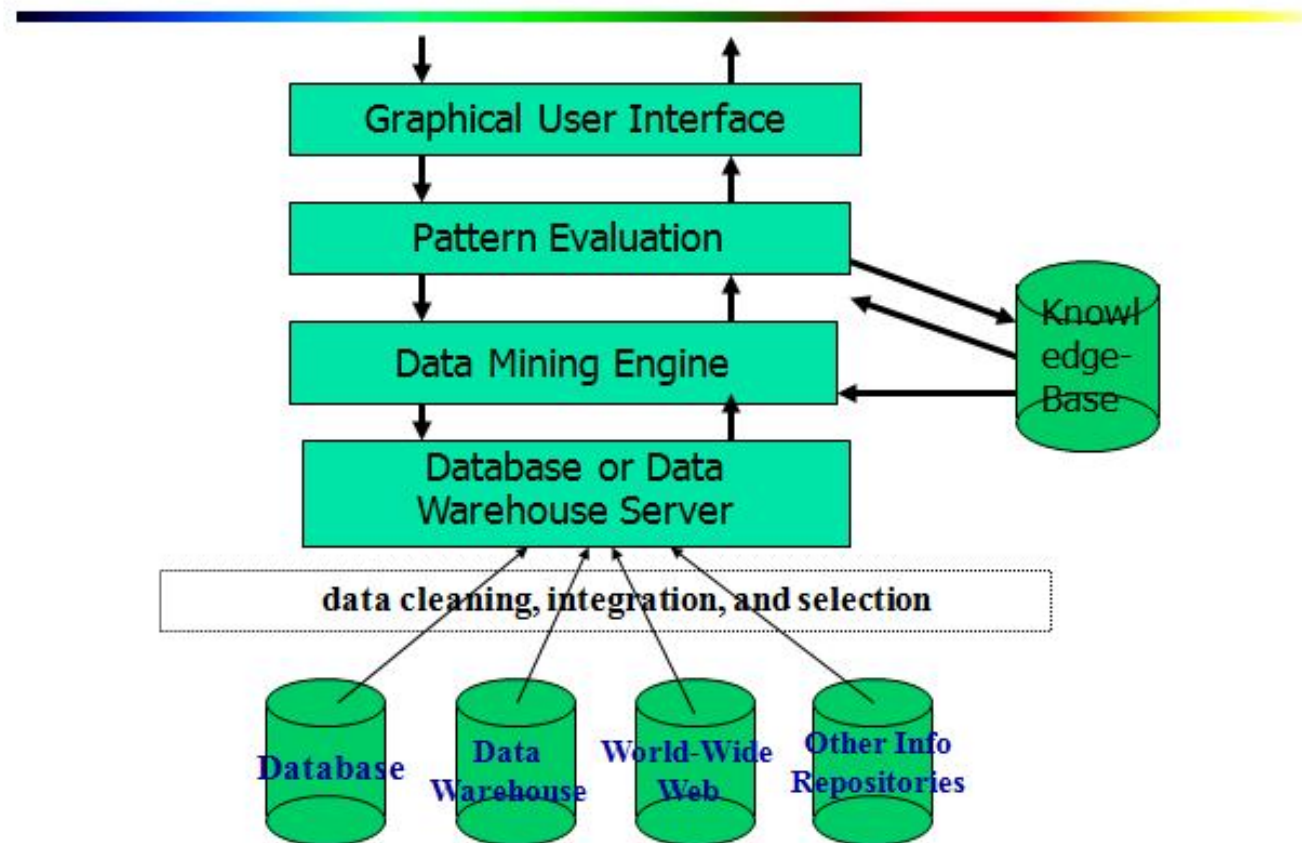
Integration of DM system with a DB/DW system

T.R.Iekhaa
AP-IT
SNSCE



Integration of DM system with a DB/DW system

Architecture: Typical Data Mining System





Integration of DM system with a DB/DW system

Possible integration schemes include;

- No coupling
- Loose coupling
- Semi tight coupling
- Tight coupling



No coupling

- Means that a DM system **will not utilize** any function of a DB OR DW system.
- It may **fetch data** from a particular source, **process data** using some data mining algorithms and then **store the mining results** in another file.



No coupling

- Such a system, though simple suffers from several drawbacks;
- ✓ DB provides a great deal of flexibility and efficiency at storing, accessing and processing data. Without a DB/DW system, a DM system may **spend a substantial amount of time** finding, collecting, cleaning and transforming data.



LOOSE COUPLING

- Means that a system will **use some facilities** of a DB or DW system, fetching data from a data repository managed by these systems, performing DM and then storing the mining results either in a file or in a DB/DW.
- Loose coupling is **better** than no coupling because it can fetch any portion of data stored in DB or DW by using query processing, indexing and other system facilities
- Advantages – flexibility and efficiency.



Semi tight coupling

- Means that besides linking a DM system to a DB/DW system, **efficient implementation** of a few essential **data mining primitives** can be provided in the DB/DW system.
- These primitives can include sorting, histogram analysis, indexing, aggregation etc



Tight coupling

- Means that a DM system is **smoothly integrated** into the DB/DW system.
- Data mining subsystem is treated as a **functional component** of an information system
- **Data mining queries and functions are optimized** (efficient/perfect) based on processing methods of a DB or DW system



Thank you