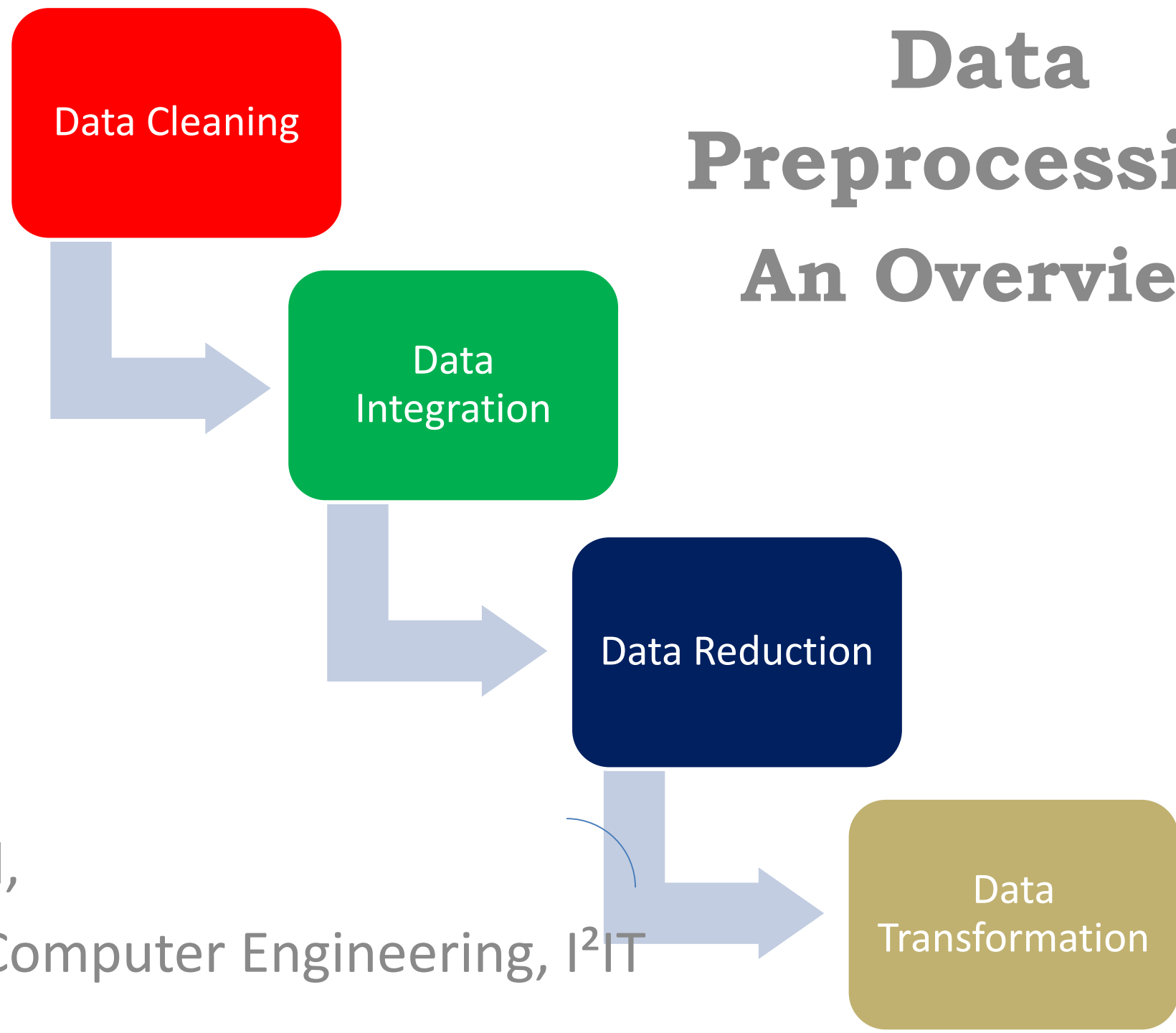


Data Preprocessing An Overview



By Sandeep Patil,
Department of Computer Engineering, I²IT

Outline

- What is Data Preprocessing ?
- Major Steps in Data Preprocessing
 - ❖ Data Cleaning
 - ❖ Data Integration
 - ❖ Data Reduction
 - ❖ Data Transformation and Data Discretization
- Conclusion

Why Data Preprocessing ?

- **Need of data preprocessing**

Some part of Data may have problems like

- Incomplete (absence of data)
- Inaccurate or noisy (other than expected values)
- Inconsistent (containing discrepancies)
- Timeliness (old version of data)
- Believability (users faith in the correctness of the data)
- Interpretability (simplicity in understanding the data)