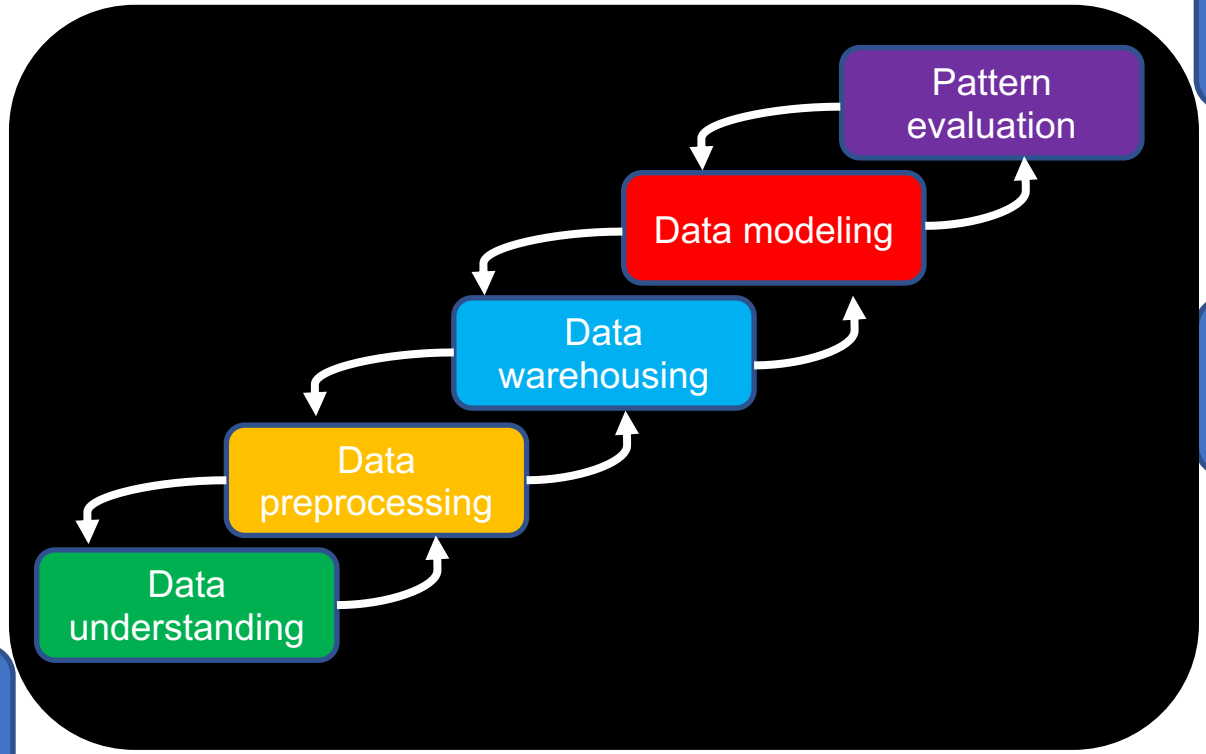
Project Checkpoint

Data Mining:
Data Mining Project
with **Dr. Qin Lv**

Master of Science in Data Science
UNIVERSITY OF COLORADO **BOULDER**

Learning objective: Design and develop real-world solutions across the full data mining pipeline.

Data Mining Pipeline



Application

Knowledge

Technique

Data

Project Progress (1)

- Obtain the data you need
- Familiarize yourself with the tools
- Data understanding
- Data preprocessing
- Data warehousing

Project Progress (2)

- ➢ Data modeling
  - Frequent patterns, classification, clustering, anomaly detection, complex data mining (text, graph, temporal)
- ➢ Pattern evaluation
  - Effectiveness, efficiency
  - Comparison, tradeoffs

Project Checkpoint

- Project status check: Are things on track?
  - progress, changes
- What to submit
  - Checkpoint slides & checkpoint report
  - Updated from proposal slides & proposal report

Checkpoint Slides

- 10-15 slides, highlight progress & changes
- Project overview
  - Project title, problem statement, related work, proposed work, evaluation, timeline
- Style: clean, large font, color, picture

Checkpoint Report

- ACM proceedings template, 3-5 pages
- Project title
- Abstract, introduction, related work, proposed work, evaluation, discussion, conclusion, references

Project Title

- ”~~Data Mining Project~~”
- Concise and informative
- Keywords to include
- (optional) acronym or codename

Abstract

- Aka: Executive Summary
- Brief summary of the project
- Usually 1-2 paragraphs

Introduction

- What is the problem?
- Why is it important?
- Limitations of existing solutions
- What is your (potential) contribution?

Related Work

- What has been done already?
- Group by category
  - E.g., tools, methods, results
- How your project builds upon prior work?
  - E.g., utilize, compare, improve, new knowledge

Proposed Work

- ➢ Datasets, tools
- ➢ Main tasks
  - E.g., statistical analysis, visualization
  - E.g., data preprocessing, warehousing
  - E.g., specific questions, patterns to explore/model

Evaluation

- Evaluation metrics
  - Effectiveness and efficiency
- Experimental setup
- Methods to compare
- Preliminary results

Discussion

- Project timeline
  - When to finish what, current status
- (Potential) challenges
- Alternative approaches/backup plan
- Changes, lessons

Conclusion

- Project summary
- (when finished) Key findings
- (when finished) Future work

Project Checkpoint

- Project status check: Are things on track?
  - progress, changes
- What to submit
  - Checkpoint slides & checkpoint report
  - Updated from proposal slides & proposal report