

Project Checkpoint Review

**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.

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 Slides: Review

- Good project overview?
- Highlight progress & changes
 - What has been accomplished so far?
 - What are the changes, if any?
- 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

Checkpoint Report: Review

- Updated/expanded sections
- Project title
- Abstract, introduction, related work, proposed work, evaluation, discussion, conclusion, references

Project Progress (1)

- Obtain the data
- Familiarize 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

Key Questions to Consider

- Is the project **on track**?
 - Progress, changes, preliminary results
- **Analytical thinking**
 - Reason about the process, methods, results
- Any **suggestions**? Any **lessons**?