**Unit 3.2 Assignment: Summaries and Abstracts**

Dennis Clement

Post University

CIS 311\_30 Technical Writing in CIS

Dr. Matthew Zullo

September 13th, 2025

**Data Mining**

This study offers a thorough analysis of the literature about the adaptation of prevalent data mining methodologies across various sectors. The authors analyze several research publications to differentiate between established methodologies (e.g., CRISP-DM, SEMMA, and KDD) and the modifications or enhancements adopted by practitioners to address practical challenges. Many initiatives don't really employ any methodology; instead, they tweak or leave out parts of methodologies. This is something that is fun to see. Some common changes include rearranging procedures, skipping steps that aren't necessary (like getting data ready or testing a model), adding steps that are important to the region, and combining steps to save time. Some of the questions that have come up are how to apply domain knowledge, how to deal with changing data, how to deal with data quality, and how to make procedures work with large amounts of data or when data is being streamed. The authors say that modifying the methodology is often essential because of time, money, the subject's intricacy, or privacy and regulatory considerations. They also say that there isn't much information on how to get better at a skill. Many publications do not clarify the distinctions that set their processes apart from commonly utilized methods. Lastly, the study says that we should make it easy for people to use more than one approach by giving them templates that can be changed to fit different needs and by looking at how changes affect the quality, repeatability, and reliability of models.

This article examines literature regarding the implementation of data mining approaches in practical contexts. The authors do a comprehensive examination of existing literature to compare conventional approaches (e.g., CRISP-DM, KDD, SEMMA) with their adapted applications in real-world projects. They discuss about how people usually adapt by leaving things out, changing the sequence of steps, merging steps, or adding their own. They also identify the challenges that cause individuals adapt, such as not understanding enough about the subject, low data quality, changing data, or not having enough resources. The breadth encompasses employing methodologies from other disciplines, examining the manner in which researchers articulate their findings, and identifying issues related to documentation, reproducibility, and methodological rigor. The purpose is to help people learn how to adapt and provide them ideas for how to make methodology frameworks more flexible or better documented.

The authors of "Adaptations of Data Mining Methodologies: A Systematic Literature Review" aim to clarify the mechanisms and justifications for the practical alterations of conventional data mining methodologies (e.g., CRISP-DM, SEMMA, KDD). They systematically analyzed existing papers, categorizing the phases that authors alter, omit, or rearrange in their methodologies, and identifying the factors that drive these modifications. Significant findings indicate that numerous projects do not strictly conform to a singular methodology; prevalent modifications involve the simplification or omission of preprocessing, evaluation, or interpretation steps; domain knowledge frequently necessitates the incorporation of additional procedures; real-time or streaming data, as well as big data contexts, demand methodological alterations for scalability; and insufficient documentation regarding the adaptation of methodologies compromises reproducibility and transparency. The authors assert that adaptation is both prevalent and significant; nonetheless, they contend that methodological studies must provide greater clarity regarding the modifications made and their rationale. They recommend developing flexible methodological templates, establishing clearer guidelines for adaptation decisions, and conducting further research on the impact of adaptations on outcomes including model accuracy, trust, and repeatability.

There are several reasons for using descriptive and informative abstracts. A descriptive abstract gives a brief overview of the article's main ideas, topics, and themes without giving away any conclusions or comprehensive findings. It helps readers figure out if they want to read the article. The informative abstract, on the other hand, tells you not only what the paper does, but also how it does it and what it found. It gives enough information that someone may understand the main points without reading the whole manuscript.

In a professional or academic setting:

When the audience wants to swiftly read a lot of papers (such during a literature review, conference proceedings, or cataloging), and judgments on deeper reading are largely based on relevance, topic, or scope rather than outcomes, use descriptive abstracts.

Use informative abstracts when decisions depend on real outcomes, such choosing methodologies, deciding whether to accept results, making policy or business decisions, or giving a summary to stakeholders who might not read the whole thing.

Informative abstracts are better for making decisions since they let you think about whether the article's conclusions are relevant, trustworthy, or helpful. When you're filtering or mapping the field, descriptive abstracts are better.

The informative abstract was harder for me because I had to decide which outcomes and methodological information to include (without making it too long) based on what I thought was most important. It is harder to make sure that anything is accurate, clear, and balanced (without putting too much weight on small aspects) than to only summarize its scope and purpose.

**References**

Adaptations of data mining methodologies: A systematic ... (n.d.). <https://peerj.com/articles/cs-267.pdf>