Notice of the Final Oral Examination
for the Degree of Master of Science

of

ALVI MAHADI

BSc (Institute of Information Technology, Jahangirnagar University)

“Conclusion Stability for Natural Language Based Mining of Design Discussions”

Department of Computer Science

Tuesday, February 2, 2021
10:00 A.M.
Remote Defence

Supervisory Committee:
Dr. Neil Ernst, Department of Computer Science, University of Victoria (Supervisor)
Dr. Daniel German, Department of Computer Science, UVic (Member)

External Examiner:
Dr. Michael Famelis, Department of Computer Science and Operations Research, Université de Montréal

Chair of Oral Examination:
Dr. Jens Bornemann, Department of Electrical and Computer Engineering, UVic

Dr. Stephen Evans, Acting Dean, Faculty of Graduate Studies
Abstract

Developer discussions range from in-person hallway chats to comment chains on bug reports. Being able to identify discussions that touch on software design would be helpful in documentation and refactoring software. Design mining is the application of machine learning techniques to correctly label a given discussion artifact, such as a pull request, as pertaining (or not) to design. In this work we demonstrate a simple example of how design mining works. We first replicate an existing state-of-the-art design mining study to show how conclusion stability is poor on different artifact types and different projects. Then we introduce two techniques—augmentation and context specificity—that greatly improve the conclusion stability and cross-project relevance of design mining. Our new approach achieves AUC-ROC of 0.88 on within dataset classification and 0.84 on the cross-dataset classification task.