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

of

UTTEJ REDDY GONDHI

Bech (Jawaharlal Nehru Technological University, 2016)

“Intra-Topic Clustering for Social Media”

Department of Electrical and Computer Engineering

Friday, August 14, 2020
10:30 A.M.
Remote Defence

Supervisory Committee:
Dr. Stephen Neville, Department of Electrical and Computer Engineering, University of Victoria
(Supervisor)
Dr. Michael McGuire, Department of Electrical and Computer Engineering, UVic (Member)

External Examiner:
Dr. Ulrike Stege, Department of Computer Science, UVic

Chair of Oral Examination:
Dr. Francis Lau, Chair, School of Health Information Science, UVic

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

With the social media platforms leading the internet in terms of user base and the average time spent, so is the amount of data that is being generated by these platforms every day. This makes social media platforms a go-to place to understand the reviews, trends, and opinions of the people. Any regular search for a popular topic would result in an abundance of information and thus it is impossible to go through these large amounts of data manually to understand the trends.

This thesis discusses techniques for the intra-topic clustering of such social media data. The social media noise increasing the redundancy of the search results are discussed and the goal is to filter the amount of redundant information an end-user must review from a regular social media search. The research proposed clustering models based on two string similarity measures, Jaccard word token, and T-Information distance. The evaluation parameters are introduced, and the models are evaluated on clustering the current and historical topics to determine which techniques were effective.