



**University  
of Victoria**

Graduate Studies

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

of

**MOHAMMADALI DEHGHAN**

BSc (Amirkabir University of Technology, 2015)

**“Predicting Likelihood of Requirement Implementation within  
the Planned Iteration”**

Department of Computer Science

Friday, May 12, 2017

1:00 P.M.

Engineering and Computer Science Building  
Room 548

Supervisory Committee:

Dr. Daniela Damian, Department of Computer Science, University of Victoria (Co-Supervisor)  
Dr. Kelly Blincoe, Department of Computer Science, UVic (Co-Supervisor)

External Examiner:

Dr. Daniela Constantinescu, Department of Mechanical Engineering, UVic

Chair of Oral Examination:

Dr. Xiaodai Dong, Department of Electrical and Computer Engineering, UVic

Dr. David Capson, Dean, Faculty of Graduate Studies

## **Abstract**

There has been a significant interest in the estimation of time and effort in fixing defects among both software practitioners and researchers over the past two decades. However, most of the focus has been on prediction of time and effort in resolving bugs, or other low level tasks, without much regard to predicting time needed to complete high-level requirements, a critical step in release planning. In this thesis, we describe a mixed-method empirical study on three large IBM projects in which we developed and evaluated a process of training a predictive model constituting a set of 29 features in nine categories in order to predict if whether or not a requirement will be completed within its planned iteration. We conducted feature engineering through iterative interviews with IBM software practitioners as well as analysis of large development and project management repositories of these three projects. Using machine learning techniques, we were able to make predictions on requirement completion time at four different stages of requirement lifetime. Using our industrial partner's interest in high precision over recall, we then adopted a cost sensitive learning method and maximized precision of predictions (ranging from 0.8 to 0.97) while maintaining an acceptable recall. We also ranked the features based on their relative importance to the optimized predictive model. We show that although satisfying predictions can be made at early stages, even on the first day of requirement creation, performance of predictions improves over time by taking advantage of requirements' progress data. Furthermore, feature importance ranking results show that although importance of features are highly dependent on project and prediction stage, there are certain features (e.g. requirement creator, time remained to the end of iteration, time since last requirement summary change and number of times requirement has been replanned for a new iteration) that emerge as important across most projects and stages, implying future worthwhile research directions for both researchers and practitioners.