Contents

Certificate iii

Declaration by the Candidate v

Copyright Transfer Certificate vii

Preface xi

Acknowledgements xv

Contents xvii

List of Figures xxiii

List of Tables xxvii

Abbreviations xxxi

Symbols xxxiii

1 Introduction 1

1.1 Link prediction . . . 3
1.2 Challenges in Link Prediction . . . 4
1.3 Motivation of the thesis . . . 5
1.4 Contribution of the thesis . . . 8
1.4.1 Feature fusion based link prediction in dynamic network . . . 8
1.4.2 Path Weight Aggregation Feature for link prediction . . . 9
1.4.3 A new cost based feature for link prediction . . . 10
1.4.4 Community Enhanced Link Prediction . . . 10
1.4.5 Projected Quantum Kernel based Link prediction . . . 11

1.5 Organization of the thesis 11

2 Literature Review **13**
2.1 Similarity-based methods 13
2.1.1 Local similarity based indices 15
2.1.2 Global similarity indices 17
2.1.3 Quasi-local indices 18
2.1.4 Clustering based Features 21
2.2 Dynamic Networks 23
2.2.1 Dynamic Datasets 25
2.2.2 Link Prediction Framework in dynamic networks 27
2.3 Machine learning classifiers in link prediction 30
2.3.1 Machine learning classifier 31
2.3.2 Ensemble learning 33
2.3.3 Performance evaluation metrics 34

3 Features Fusion based Link Prediction in Dynamic Networks **37**
3.1 Introduction 38
3.2 Proposed work 41
3.2.1 Proposed Feature generation model Algorithm 42
3.3 Result Analysis 43
3.3.1 Performance of LGQ model and its variations with Neural Network (NN) 43
3.3.2 Performance of LGQ and its variations with XGBoost 47
3.3.3 Comparison of LGQ model with state-of-the-art methods 48
3.4 Conclusion 51

4 Path Weight Aggregation Feature for Link Prediction in Dynamic Networks (PWAF) **53**
4.1 Introduction 54
4.2 Proposed work 57
4.2.1 Path Weight Aggregation Feature $(PWAF)$ for link prediction . . 57
4.2.2 Analysis for selecting $\sqrt{N}$ 59
4.2.3 Proposed PWAF Feature Generation Algorithm 59
4.3 Result Analysis 63
4.3.1 Performance of PWAF model and its machine learning variations-XGBoost (XGB) 67
4.3.2 Performance of PWAF model and its machine learning variations-Random Forest Classifier (RFC) 68
4.3.3 Performance of PWAF model and its machine learning variations-Linear Discriminant Analysis (LDA) 68

4.3.4 Performance comparison of PWAF model machine learning variations . . . 69
4.3.5 Comparison of PWAF variations with state-of-the-art methods . . 71
4.4 Conclusion . . . 73

5 A New Cost Based Feature for Link Prediction in Dynamic Networks (CFLP) 75
5.1 Introduction . . . 76
5.2 Proposed Method . . . 78
5.2.1 A New Cost Based Feature for Link Prediction(CFLP) algorithm . 78
5.2.2 Algorithm for training and testing of CFLP model . . . 79
5.2.3 Feature score used in CFLP . . . 83
5.3 Result Analysis . . . 85
5.3.1 Performance comparison of CFLP model with different similarity methods on Neural Network (NN) . . . 85
5.3.2 Performance comparison of CFLP model with different similarity methods on Logistic Regression (LR) . . . 86
5.3.3 Performance comparison of CFLP model with different similarity methods with Extreme Gradient boosting XGBoost (XGB) . . . 89
5.3.4 Performance comparison of CFLP model with different similarity methods with Random Forest Classifier (RFC) . . . 91
5.3.5 Performance comparison of CFLP model with different similarity methods Linear Discriminant Analysis (LDA) . . . 91
5.3.6 Comparison of CFLP with state-of-the-art methods after optimization . . . 94
5.4 Conclusion . . . 97

6 Community Enhanced Link Prediction in Dynamic Networks 99
6.1 Introduction . . . 99
6.1.1 Community Detection . . . 102
6.1.2 Community-based Feature Generation Methods . . . 103
6.1.3 Classification Models . . . 106
6.2 Proposed work . . . 108
6.2.1 Topological Feature Generation . . . 109
6.2.2 Community-based Feature Generation . . . 109
6.2.3 Feature Set Engineering and Link Prediction . . . 112
6.2.4 Feature Reduction . . . 113
6.2.5 Algorithm Description . . . 114
6.2.6 Demonstration with Example . . . 116
6.3 Result Analysis . . . 118
6.3.1 Performance comparison of $COMMLP - FULL$ with individual feature based methods using Neural Network model . . . 119

6.3.2 Performance comparison of *COMMLP – FULL* with individual feature based methods using XGBoost model 121
6.3.3 Performance comparison of *COMMLP – FULL* with individual feature based methods using Linear Discriminant Analysis model 124
6.3.4 Performance comparison of *COMMLP – FULL* with individual feature based methods using Random Forest Classifier model . . 124
6.3.5 Comparison of Individual and Community Information based Link Prediction Features 126
6.3.6 Comparison of *COMMLP-DYN* with State-of-the-Art Methods after Optimization 129
6.4 Conclusion 132

7 Projected Quantum Kernel based Link Prediction in Dynamic Networks (PQKLP) 133
7.1 Introduction 134
7.1.1 Quantum Computation 137
7.1.2 Projected Quantum Kernels (PQK) 139
7.2 Proposed Method 145
7.2.1 Algorithms used in the proposed framework 145
7.2.2 Feature generation algorithm for proposed PQKLP approach . . . 148
7.3 Result Analysis 149
7.3.1 Performance comparison and analysis of PQKLP model with its component individual link prediction methods using Neural Network (NN) 149
7.3.2 Performance comparison and analysis of PQKLP model with its component individual link prediction methods using Logistic Regression (LR) 151
7.3.3 Performance comparison and analysis of PQKLP model with its component individual link prediction methods using Linear Discriminant Analysis (LDA) 153
7.3.4 Performance comparison and analysis of PQKLP model with its component individual link prediction methods using XGBoost (XGB) 153
7.3.5 Performance comparison and analysis of PQKLP model with its component individual link prediction methods using Random Forest Classifier (RFC) 154
7.3.6 Performance comparison and analysis of PQKLP model with its component individual link prediction methods using Gaussian Naive Baiyes classifier (GNB) 156
7.3.7 Performance comparison and analysis of classical machine learning model with proposed PQKLP-based machine learning model with same feature set 157

7.3.8 Performance comparison and analysis of Projected Quantum Kernel based Link Prediction (PQKLP) model with various state-of-the art algorithms 159
7.4 Conclusion 160

8 Conclusion and future directions **163**
8.1 Summary of Contributions 163
8.2 Scope for Future Work 166

Bibliography **169**

A List of Publications **193**