CONTENTS

Acknowledgement i

Abstract ii

List of Figures v

List of Tables xii

List of Abbreviations xvi

1. INTRODUCTION 1

1.1 Data Mining 2

1.2 Overview of Spinocerebellar Ataxia Disorder 11

1.3 Docking and Binding Affinity 26

1.4 Need for the Proposed Research 35

1.5 Machine Learning and Deep Learning 36

1.5.1 Regression Algorithms 36

1.5.2 Deep Neural Networks 42

1.6 Problem Statement and Objectives of the Research 48

1.7 Organization of the Thesis 49

2 REVIEW OF LITERATURE 51

2.1 Traditional Approaches 51

2.2 Computational Methods 53

3 PROBLEM MODELLING 63

3.1 Overall Framework of Binding Affinity Prediction Model 63

3.2 Development of Corpuses 65

3.3 Design of Features And Datasets 78

3.4 Training And Testing 82

4 BINDING AFFINITY PREDICTION MODEL USING PROTEIN-LIGAND DOCKING AND REGRESSION TECHNIQUES 86

4.1 Protein-Ligand Docking Based Binding Affinity Predictive Models using Supervised Learning 86

4.2 Experiment and Results 97

5 BINDING AFFINITY PREDICTION USING PROTEIN-MUTATED-LIGAND DOCKING AND REGRESSION TECHNIQUES 103

5.1 Protein-Mutated-Ligand Docking Based Binding Affinity Predictive Models using Supervised Learning 103

5.2 Experiment and Results 117

6 BINDING AFFINITY PREDICTION USING PROTEIN-PROTEIN INTERACTIONS AND REGRESSION TECHNIQUES 126

6.1 Protein-Protein Interaction Based Binding Affinity Predictive Models using Supervised Learning 126

6.2 Experiment and Results 134

7 BINDING AFFINITY PREDICTION USING DEEP NEURAL NETWORK ARCHITECTURES 143

7.1 Binding Affinity Prediction Models using Sequential Deep Neural Network 144

7.2 Binding Affinity Prediction Models using Functional Deep Neural Network 162

7.3 Binding Affinity Prediction Models using Deep Neural Network with Customized Layers 178

7.4 Comparative Analysis 195

8. CONCLUSION 201

REFERENCES 204

List of Publications 213

Appendix

A. Datasets 215

B. Coding 225

C. Screenshots 233