TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER NO.</th>
<th>TITLE</th>
<th>PAGE NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td></td>
<td>LIST OF TABLES</td>
<td>xiii</td>
</tr>
<tr>
<td></td>
<td>LIST OF FIGURES</td>
<td>xv</td>
</tr>
<tr>
<td></td>
<td>LIST OF SYMBOLS AND ABBREVIATIONS</td>
<td>xvii</td>
</tr>
<tr>
<td>1</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>SOFTWARE ENGINEERING</td>
<td>2</td>
</tr>
<tr>
<td>1.2.1</td>
<td>Software Development Life Cycle (SDLC)</td>
<td>3</td>
</tr>
<tr>
<td>1.3</td>
<td>INTRODUCTION TO REQUIREMENT ENGINEERING</td>
<td>5</td>
</tr>
<tr>
<td>1.3.1</td>
<td>The Importance of Requirements Engineering</td>
<td>7</td>
</tr>
<tr>
<td>1.3.2</td>
<td>Types of Requirements</td>
<td>8</td>
</tr>
<tr>
<td>1.3.3</td>
<td>Requirements and Quality</td>
<td>9</td>
</tr>
<tr>
<td>1.3.4</td>
<td>Requirements and Modeling</td>
<td>10</td>
</tr>
<tr>
<td>1.3.5</td>
<td>Requirements and Testing</td>
<td>11</td>
</tr>
<tr>
<td>1.3.6</td>
<td>Requirements in the Problem and Solution Domains</td>
<td>13</td>
</tr>
<tr>
<td>1.3.7</td>
<td>Requirements Traceability</td>
<td>15</td>
</tr>
<tr>
<td>1.4</td>
<td>TRACING DEFECTS TO REQUIREMENTS</td>
<td>19</td>
</tr>
<tr>
<td>1.5</td>
<td>FUNCTIONAL REQUIREMENTS AND NON FUNCTIONAL REQUIREMENTS</td>
<td>21</td>
</tr>
<tr>
<td>1.5.1</td>
<td>Why Non Functional Requirements?</td>
<td>23</td>
</tr>
<tr>
<td>CHAPTER NO.</td>
<td>TITLE</td>
<td>PAGE NO.</td>
</tr>
<tr>
<td>-------------</td>
<td>-------</td>
<td>----------</td>
</tr>
<tr>
<td>1.5.2</td>
<td>Approaches for dealing with Non Functional Requirements</td>
<td>23</td>
</tr>
<tr>
<td>1.5.3</td>
<td>Dealing with Non Functional Requirements</td>
<td>25</td>
</tr>
<tr>
<td>1.6</td>
<td>MOTIVATION</td>
<td>26</td>
</tr>
<tr>
<td>1.7</td>
<td>OBJECTIVES OF THE THESIS</td>
<td>26</td>
</tr>
<tr>
<td>1.8</td>
<td>THESIS OVERVIEW</td>
<td>27</td>
</tr>
<tr>
<td>2</td>
<td>LITERATURE REVIEW</td>
<td>29</td>
</tr>
<tr>
<td>2.1</td>
<td>LITERATURE REVIEW</td>
<td>29</td>
</tr>
<tr>
<td>2.2</td>
<td>LITERATURE REVIEW ON REQUIREMENTS ENGINEERING</td>
<td>29</td>
</tr>
<tr>
<td>2.3</td>
<td>LITERATURE REVIEW ON DEFECT TRACING TO REQUIREMENTS</td>
<td>34</td>
</tr>
<tr>
<td>2.4</td>
<td>LITERATURE REVIEW ON INFORMATION RETRIEVAL FOR REQUIREMENTS ENGINEERING</td>
<td>36</td>
</tr>
<tr>
<td>2.5</td>
<td>REVIEW ON NON FUNCTIONAL REQUIREMENTS (NFR)</td>
<td>37</td>
</tr>
<tr>
<td>3</td>
<td>INVESTIGATION OF INFORMATION RETRIEVAL TECHNIQUES FOR IDENTIFYING NON FUNCTIONAL REQUIREMENTS</td>
<td>47</td>
</tr>
<tr>
<td>3.1</td>
<td>INTRODUCTION</td>
<td>47</td>
</tr>
<tr>
<td>3.2</td>
<td>FUNCTIONAL AND NON FUNCTIONAL REQUIREMENTS</td>
<td>48</td>
</tr>
</tbody>
</table>
CHAPTER NO. | TITLE | PAGE NO.
--- | --- | ---
3.3 | PROPOSED METHODOLOGY | 50
3.3.1 | Dataset | 50
3.3.2 | Term Frequency | 52
3.3.3 | Bagging Methods | 53
3.3.4 | Boosting Methods | 55
3.4 | RESULTS AND DISCUSSION | 57
3.5 | SUMMARY | 66

4 | IDENTIFYING NON FUNCTIONAL REQUIREMENTS USING SOFT COMPUTING TECHNIQUES | 67
4.1 | INTRODUCTION | 67
4.2 | METHODOLOGY | 68
4.2.1 | Feed Forward Neural Network (FFNN) | 68
4.2.1.1 | Training Feed Forward Neural Network | 70
4.2.1.2 | Selecting the number of hidden layers | 70
4.2.1.3 | Deciding how many neurons to use in the hidden layers | 71
4.2.1.4 | Finding a globally optimal solution | 71
4.2.2 | Recurrent Neural Network (RNN) | 72
4.2.3 | Genetic Algorithm (GA) | 75
4.3 | METHODOLOGY | 77
4.3.1 | Modified FFNN | 77
4.3.2 | Modified RNN | 78
4.4 | RESULTS AND DISCUSSION | 81
5 IMPROVED FEATURE EXTRACTION
AND ENHANCED NEURAL NETWORK
FOR RETRIEVAL OF NON
FUNCTIONAL REQUIREMENTS

5.1 INTRODUCTION 97
5.2 METHODOLOGY 98
  5.2.1 Proposed Feature Extraction Technique - NFR
      Repository Feature Extraction (NFR-RFE) 98
  5.2.2 Proposed Hybrid Genetic Algorithm
      Continuous Ant Colony Optimization
      (HGACACO) 101
      5.2.2.1 Initialization 102
      5.2.2.2 Enhancement 102
      5.2.2.2.1 Ant Based Solution
          Construction 104
      5.2.2.2.2 Pheromone Update 104
      5.2.2.3 Crossover 105
      5.2.2.4 Mutation 106
5.3 RESULTS AND DISCUSSION 106
<table>
<thead>
<tr>
<th>CHAPTER NO.</th>
<th>TITLE</th>
<th>PAGE NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3.1</td>
<td>Results of the First Set of Experiments</td>
<td>106</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Results of the Second Set of Experiments</td>
<td>109</td>
</tr>
<tr>
<td>5.4</td>
<td>SUMMARY</td>
<td>114</td>
</tr>
<tr>
<td>6</td>
<td>CONCLUSION AND FUTURE WORK</td>
<td>115</td>
</tr>
<tr>
<td>6.1</td>
<td>CONCLUSION</td>
<td>115</td>
</tr>
<tr>
<td>6.2</td>
<td>FUTURE WORK</td>
<td>117</td>
</tr>
<tr>
<td>APPENDIX 1</td>
<td></td>
<td>118</td>
</tr>
<tr>
<td>APPENDIX 2</td>
<td></td>
<td>119</td>
</tr>
<tr>
<td>APPENDIX 3</td>
<td></td>
<td>120</td>
</tr>
<tr>
<td>APPENDIX 4</td>
<td></td>
<td>121</td>
</tr>
<tr>
<td>APPENDIX 5</td>
<td></td>
<td>122</td>
</tr>
<tr>
<td>APPENDIX 6</td>
<td></td>
<td>123</td>
</tr>
<tr>
<td>REFERENCES</td>
<td></td>
<td>124</td>
</tr>
<tr>
<td>LIST OF PUBLICATIONS</td>
<td></td>
<td>133</td>
</tr>
<tr>
<td>CURRICULUM VITAE</td>
<td></td>
<td>134</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE NO.</th>
<th>TITLE</th>
<th>PAGE NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Problem and solution spaces</td>
<td>13</td>
</tr>
<tr>
<td>1.2</td>
<td>Types of traceability analysis</td>
<td>18</td>
</tr>
<tr>
<td>3.1</td>
<td>Distribution functional and non functional requirements in the collection</td>
<td>52</td>
</tr>
<tr>
<td>3.2</td>
<td>Classification and RMSE of the technique under consideration</td>
<td>58</td>
</tr>
<tr>
<td>3.3</td>
<td>The precision and recall for the NFR class using bagging with reptree</td>
<td>60</td>
</tr>
<tr>
<td>3.4</td>
<td>The precision and recall for the NFR class using bagging with random forest</td>
<td>61</td>
</tr>
<tr>
<td>3.5</td>
<td>The precision and recall for the NFR Class using logiboost with reptree</td>
<td>61</td>
</tr>
<tr>
<td>3.6</td>
<td>The precision and recall for the NFR Class using logiboost with decision stump</td>
<td>62</td>
</tr>
<tr>
<td>4.1</td>
<td>Architecture of the neural network</td>
<td>80</td>
</tr>
<tr>
<td>4.2</td>
<td>The proposed genetic algorithm parameters</td>
<td>81</td>
</tr>
<tr>
<td>4.3</td>
<td>Mean Squared Error obtained for MLP with different activation functions in hidden and output layer</td>
<td>82</td>
</tr>
<tr>
<td>4.4</td>
<td>The precision, recall and F measure</td>
<td>85</td>
</tr>
<tr>
<td>4.5</td>
<td>Mean Squared Error obtained for GFFNN with Tanh activation function and sigmoid activation</td>
<td>86</td>
</tr>
<tr>
<td>4.6</td>
<td>The precision, recall and F measure</td>
<td>88</td>
</tr>
<tr>
<td>4.7</td>
<td>The best fitness obtained through 20 generations</td>
<td>89</td>
</tr>
<tr>
<td>4.8</td>
<td>Precision, recall and F measure of proposed HLGO-GFFFN</td>
<td>91</td>
</tr>
<tr>
<td>4.9</td>
<td>Precision, recall and F measure for HLGO-RNN</td>
<td>91</td>
</tr>
<tr>
<td>4.10</td>
<td>Classification Accuracy obtained from various techniques</td>
<td>92</td>
</tr>
<tr>
<td>4.11</td>
<td>Precision, recall and F measure for all Techniques</td>
<td>93</td>
</tr>
<tr>
<td>TABLE NO.</td>
<td>TITLE</td>
<td>PAGE NO.</td>
</tr>
<tr>
<td>-----------</td>
<td>----------------------------------------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>4.12</td>
<td>Classification accuracy compared with other techniques in literature</td>
<td>94</td>
</tr>
<tr>
<td>4.13</td>
<td>Precision of proposed technique and other techniques available in literature</td>
<td>95</td>
</tr>
<tr>
<td>5.1</td>
<td>Class Label and Seeds Used in Building the NFR Repository</td>
<td>100</td>
</tr>
<tr>
<td>5.2</td>
<td>Classification Accuracy for different Techniques</td>
<td>107</td>
</tr>
<tr>
<td>5.3</td>
<td>Precision, recall and F measure of proposed HLGO-GFFFN</td>
<td>108</td>
</tr>
<tr>
<td>5.4</td>
<td>Precision, recall and F measure for HLGO GFNN with proposed feature extraction technique</td>
<td>108</td>
</tr>
<tr>
<td>5.5</td>
<td>Precision, recall and F measure for HGACACO GFFNN classifier with proposed feature extraction technique</td>
<td>109</td>
</tr>
<tr>
<td>5.6</td>
<td>Classification Accuracy for different Techniques</td>
<td>110</td>
</tr>
<tr>
<td>5.7</td>
<td>Precision, recall and F measure for Bagging with Random forest and resampling and existing feature extraction</td>
<td>111</td>
</tr>
<tr>
<td>5.8</td>
<td>Precision, recall and F measure for Bagging with Random forest and resampling and proposed feature extraction</td>
<td>111</td>
</tr>
<tr>
<td>5.9</td>
<td>Precision, recall and F measure for HLGO GFFNN classifier with existing feature extraction technique</td>
<td>112</td>
</tr>
<tr>
<td>5.10</td>
<td>Precision, recall and F measure for HLGO GFNN classifier with proposed feature extraction technique</td>
<td>112</td>
</tr>
<tr>
<td>5.11</td>
<td>Precision, recall and F measure for HGACACO GFFNN classifier with existing feature extraction technique</td>
<td>113</td>
</tr>
<tr>
<td>5.12</td>
<td>Precision, recall and F measure for HGACACO GFFNN classifier with proposed feature extraction technique</td>
<td>113</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE NO.</th>
<th>TITLE</th>
<th>PAGE NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Qualification strategy and the V - model</td>
<td>12</td>
</tr>
<tr>
<td>3.1</td>
<td>Steps to capture non functional requirements</td>
<td>49</td>
</tr>
<tr>
<td>3.2</td>
<td>Distribution of NFRs by class in the collection of documents in data set</td>
<td>51</td>
</tr>
<tr>
<td>3.3</td>
<td>The classification accuracy of various methods</td>
<td>59</td>
</tr>
<tr>
<td>3.4</td>
<td>The RMSE obtained for the various techniques used</td>
<td>59</td>
</tr>
<tr>
<td>3.5</td>
<td>Precision using different techniques</td>
<td>63</td>
</tr>
<tr>
<td>3.6</td>
<td>Recall using different techniques</td>
<td>63</td>
</tr>
<tr>
<td>3.7</td>
<td>F measure using different techniques</td>
<td>64</td>
</tr>
<tr>
<td>3.8</td>
<td>The weighted average of precision and recall for all the techniques used</td>
<td>65</td>
</tr>
<tr>
<td>4.1</td>
<td>The multi-layer perceptron network</td>
<td>69</td>
</tr>
<tr>
<td>4.2</td>
<td>Recurrent Neural Network</td>
<td>73</td>
</tr>
<tr>
<td>4.3</td>
<td>Outline of genetic algorithm</td>
<td>76</td>
</tr>
<tr>
<td>4.4</td>
<td>Generalized Feed Forward Neural Network (GFFNN)</td>
<td>77</td>
</tr>
<tr>
<td>4.5</td>
<td>Output for steepness Parameter x=1.75</td>
<td>80</td>
</tr>
<tr>
<td>4.6</td>
<td>The average minimum Mean Squared Error for various Momentum</td>
<td>83</td>
</tr>
<tr>
<td>4.7</td>
<td>The average minimum Mean Squared Error for various learning rate for the MLP</td>
<td>83</td>
</tr>
<tr>
<td>4.8</td>
<td>Plot of average MSE</td>
<td>84</td>
</tr>
<tr>
<td>4.9</td>
<td>The average minimum mean square error for various momentum for the GFFNN</td>
<td>87</td>
</tr>
<tr>
<td>4.10</td>
<td>The average minimum Mean Squared Error for various learning rate for the GFFNN</td>
<td>87</td>
</tr>
<tr>
<td>4.11</td>
<td>The best fitness obtained</td>
<td>90</td>
</tr>
<tr>
<td>FIGURE NO.</td>
<td>TITLE</td>
<td>PAGE NO.</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>4.12</td>
<td>Classification Accuracy obtained</td>
<td>92</td>
</tr>
<tr>
<td>4.13</td>
<td>Precision and Recall</td>
<td>93</td>
</tr>
<tr>
<td>5.1</td>
<td>Block diagram of the proposed architecture</td>
<td>102</td>
</tr>
<tr>
<td>5.2</td>
<td>The best fitness obtained through 20 generations</td>
<td>107</td>
</tr>
<tr>
<td>5.3</td>
<td>Classification Accuracy for different Techniques</td>
<td>110</td>
</tr>
</tbody>
</table>
# LIST OF SYMBOLS AND ABBREVIATIONS

## Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>e</td>
<td>2.713</td>
</tr>
<tr>
<td>π</td>
<td>3.14</td>
</tr>
<tr>
<td>Ø</td>
<td>Data mining algorithm</td>
</tr>
<tr>
<td>t</td>
<td>Training data set</td>
</tr>
<tr>
<td>T</td>
<td>Training set</td>
</tr>
<tr>
<td>w</td>
<td>weight</td>
</tr>
<tr>
<td>η</td>
<td>Weighting function</td>
</tr>
<tr>
<td>α</td>
<td>Positive parameters</td>
</tr>
<tr>
<td>β</td>
<td>Positive parameters</td>
</tr>
</tbody>
</table>

## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CASE</td>
<td>Computer-Aided Software Engineering</td>
</tr>
<tr>
<td>COTS</td>
<td>Commercial Off –the- Shelf</td>
</tr>
<tr>
<td>DERAFT</td>
<td>Distributed Embedded Real-time Aspects Framework</td>
</tr>
<tr>
<td>DERTS</td>
<td>Distributed Embedded Real-time Systems</td>
</tr>
<tr>
<td>FR</td>
<td>Functional Requirements</td>
</tr>
<tr>
<td>GA</td>
<td>Genetic Algorithms</td>
</tr>
<tr>
<td>GFFNN</td>
<td>Genetic Feed Forward Neural Network</td>
</tr>
<tr>
<td>HLGO-</td>
<td>Hidden Layer Genetic Optimized Genetic Feed</td>
</tr>
<tr>
<td>GFFNN</td>
<td>Forward Neural Network</td>
</tr>
<tr>
<td>HLGO-RNN</td>
<td>Hidden Layer Genetic Optimized Recurrent Neural Networks</td>
</tr>
<tr>
<td>IR</td>
<td>Information Retrieval</td>
</tr>
<tr>
<td>KAOS</td>
<td>Knowledge Acquisition in Automated Specification</td>
</tr>
<tr>
<td>MLGFFNN</td>
<td>Multi Layer Genetic Feed Forward Neural Network</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>MLP</td>
<td>Multi Layer Perceptron</td>
</tr>
<tr>
<td>NFR</td>
<td>Non Functional Requirements</td>
</tr>
<tr>
<td>NFRE</td>
<td>Non Functional Requirements Engineering</td>
</tr>
<tr>
<td>OID</td>
<td>Object Interaction Diagram</td>
</tr>
<tr>
<td>RE</td>
<td>Requirements Engineering</td>
</tr>
<tr>
<td>REFSQ</td>
<td>Requirements Engineering Foundation for Software Quality</td>
</tr>
<tr>
<td>REGPG</td>
<td>Requirements Engineering Good Practice Guide</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root Mean Squared Error</td>
</tr>
<tr>
<td>RNN</td>
<td>Recurrent Neural Network</td>
</tr>
<tr>
<td>SDLC</td>
<td>Software Development Life Cycle</td>
</tr>
<tr>
<td>SPI</td>
<td>Software Process Improvement</td>
</tr>
<tr>
<td>SREM</td>
<td>Software Requirements Engineering Method</td>
</tr>
<tr>
<td>SRS</td>
<td>Software Requirement Specification</td>
</tr>
<tr>
<td>SVD</td>
<td>Singular Value Decomposition</td>
</tr>
<tr>
<td>TDF</td>
<td>Term Document Frequency</td>
</tr>
<tr>
<td>UAV</td>
<td>Unmanned Air Vehicle</td>
</tr>
<tr>
<td>UML</td>
<td>Unified Modeling Language</td>
</tr>
<tr>
<td>CL</td>
<td>Class Label</td>
</tr>
<tr>
<td>DCL</td>
<td>Dominant Class Label</td>
</tr>
<tr>
<td>CLT</td>
<td>Class Label Tendency</td>
</tr>
<tr>
<td>DM</td>
<td>Discriminative Measure</td>
</tr>
<tr>
<td>AS</td>
<td>Ant System</td>
</tr>
<tr>
<td>ACO</td>
<td>Ant Colony Optimization</td>
</tr>
<tr>
<td>RFE</td>
<td>Repository Feature Extraction</td>
</tr>
<tr>
<td>HGACACO</td>
<td>Hybrid Genetic Algorithm Continuous Ant Colony Optimization</td>
</tr>
<tr>
<td>ACS</td>
<td>Ant Colony System</td>
</tr>
<tr>
<td>M.E</td>
<td>Master of Engineering</td>
</tr>
</tbody>
</table>