

Table of Contents

Chapter 1: Introduction	1
1.1 Introduction	1
1.2 Problem Description	3
1.3 Research Hypothesis	5
1.4 Motivation	5
1.5 Motivation of Using Data Mining on Healthcare Data	6
1.6 Medical Billing Domain	7
1.7 Technology Domain	8
1.7.1 Learning Apprentice Systems	9
1.7.2 Knowledge Representation	10
1.7.3 Rule Based Systems	10
1.8 Research Contribution	15
1.9 Thesis Review	16
1.10 Chapter Summary	18
Chapter 2: Literature Review	19
2.1 Introduction	19
2.2 Rule Based Learning Apprentice Systems	20
2.3 Apprenticeship Learning Techniques	22
2.3.1 Explanation Based Learning	23
2.3.2 Similarity Based Learning	24
2.3.3 Integrated Learning	25
2.3.4 Data Mining Based Learning Methodology	25
2.4 Architecture of POEMS	26
2.5 Data vs. Knowledge	30
2.6 Data Mining	31
2.6.1 Mining of Rules	33
2.6.2 Association Rule Mining	34
2.6.3 Classification Rule Mining	36

2.6.4 Production Rule Mining	37
2.7 Healthcare IT	38
2.8 Data Mining In Healthcare	39
2.9 Claim Scrubbing for Medical Billing Compliance	41
2.10 Chapter Summary	42
Chapter 3: Data Mining Driven Learning Apprentice System for Medical Billing Compliance	44
3.1 Introduction	44
3.2 Medical Billing Process	44
3.3 Architecture of Data Mining Driven Learning Apprentice System	47
3.4 Components of the System	48
3.4.1 Knowledge Base	49
3.4.2 Knowledge Editor	49
3.4.3 Rule Engine	49
3.4.4 Data Warehouse of the System	50
3.4.5 Data Mining Based Learning Module	51
3.5 External Components	52
3.5.1 Medical Billing Software	52
3.5.2 Company's Website	52
3.5.3 EMR and Synchronization Server	52
3.5.4 Import Files (HL7, Amazing Charts)	52
3.5.5 Servers of New Practices	53
3.6 Domain Users	53
3.6.1 Medical Providers	53
3.6.2 Practices' Office Staff	54
3.6.3 Billing Executives	54
3.6.4 Domain Experts	54
3.6.5 Knowledge Engineers	54
3.6.6 RBS Developers	55
3.7 Study Flow	55
3.8 Performance Measurement	56

3.8.1 Processing Time Saved _____	57
3.8.2 Aging of Claims with Same Errors _____	57
3.8.3 Average Development Time of Checks and Rules _____	58
3.9 Implementation Tools of the Proposed System _____	58
3.10 Chapter Summary _____	59
Chapter 4: Data Warehouse Design and Algorithm for Production Rule Mining _____	61
4.1 Introduction _____	61
4.2 Need of the Data Warehouse _____	61
4.3 Proposed Design of the Data Warehouse _____	62
4.3.1 Fact Tables _____	64
4.3.2 Dimension Tables _____	65
4.3.3 Data Import Module _____	67
4.4 Production Rule Mining _____	67
4.5 Master Algorithm (A1) for Production Rule Mining; _____	70
4.6 Record Couple Matching Sub-algorithm (A2) _____	72
4.7 String Generalization Sub-algorithm (B1): _____	74
4.8 Condition Mining Sub-algorithm (A3): _____	75
4.9 Rule Generation Sub-algorithm (A4) _____	75
4.10 Production Rule Mining From Medical Billing Data _____	77
4.10.1 Algorithm Objects in Medical Billing Domain _____	77
4.10.2 Testing of Production Rule Mining Algorithm _____	78
4.11 Time Complexity Analysis of the Proposed Algorithm _____	87
4.12 Applying Apriori Algorithm _____	88
4.12.1 Original Apriori Algorithm _____	89
4.12.2 Apriori-Modified Algorithm _____	91
4.13 Modifying Inductive Learning Algorithm for Mining of Production Rules _____	94
4.13.1 Formulation of ILA-modified _____	95
4.13.2 The Inductive Learning Algorithm Modified (ILA-modified) _____	95
4.14 Performance Comparison of the Proposed Algorithm, Apriori-modified and ILA-modified _____	97

4.15 Applying the Proposed Algorithm on Real Life Data	99
4.16 Accuracy of the Proposed Algorithm	102
4.17 Validation of Extracted Rules	103
4.18 Chapter Summary	104
Chapter 5: SQL Based Knowledge Representation and Knowledge Editor	105
5.1 Introduction	105
5.2 Knowledge Representation	105
5.3 Representing Production Rules as Dynamic SQL Queries	106
5.4 Knowledge Base of the System	107
5.4.1 Logical Variables	107
5.4.2 Meta-Rules	108
5.4.3 Rules	108
5.4.4 Rule Action Procedures	111
5.4.5 Rule Application Log	111
5.4.6 Rule Application Log History	112
5.5 Rule Development Process	113
5.6 Knowledge Editors	114
5.7 Knowledge Editor of the RBS	115
5.7.1 Knowledge Editor	116
5.7.2 Meta-Rule Editing	121
5.7.3 Logical Variable Editor	121
5.7.4 Rule Debugger	122
5.7.5 Rule Management Module	123
5.7.6 Knowledge Base Security Module	125
5.7.7 RBS Performance Monitor	125
5.7.8 Rule Learner Module	126
5.7.9 RBS Dynamic Reports	127
5.7.10 Rule Serialization	127
5.8 Execution of Rules	128
5.9 Chapter Summary	129

Chapter 6: Rule Based Engine Design _____	130
6.1 Introduction _____	130
6.2 Rule Based Engine using SQL _____	131
6.3 Design of Rule Based Engine _____	131
6.3.1 Initial Engine Design _____	131
6.3.2 Sample Rules _____	137
6.4 Weaknesses of Initial Rule Engine Design _____	140
6.5 Multi-Pass Rule Engine Design _____	142
6.5.1 Working Memory (WM) _____	144
6.5.2 Concurrency _____	144
6.6 Enhanced Rule Engine Design _____	144
6.7 RBE vs. Claim Scrubber (implemented in SQL) _____	146
6.8 Rule Based Engine vs. Conventional Software _____	149
6.9 Performance Evaluation of the Overall System _____	151
6.10 Chapter Summary _____	151
Chapter 7: Results and Conclusion _____	153
7.1 Introduction _____	153
7.2 Evaluation of Rule Based Expert Systems _____	153
7.3 RBS Performance Analysis _____	155
7.3.1 Rule Development _____	157
7.3.2 Claim Scrubbing _____	158
7.4 Measurement of Benefits _____	160
7.4.1 Gain 1 (Development Time Saved) _____	161
7.4.2 Gain 2 (Processing Time and Claim Aging) _____	169
7.5 Indirect Measures of Performance Analysis _____	178
7.5.1 Claim Rejection Report _____	179
7.5.2 Claim Blocking _____	180
7.6 RBS Cost Analysis _____	183
7.7 Return on Investment (ROI) _____	184
7.8 Discussion _____	185

7.8.1 The Basic Philosophy of Rule Based Systems _____	186
7.8.2 The Basic Philosophy of Data Mining _____	189
7.9 Limitations of the Research Work _____	190
7.9.1 Suitable Application Domains _____	190
7.9.2 Mining of ‘How’ Actions are performed _____	190
7.9.3 One Value to Process _____	191
7.9.4 Burden on Database Server _____	191
7.9.5 History Related Rules _____	191
7.10 Contribution Summary of the Research Work _____	192
7.10.1 Production Rule Mining Algorithm _____	192
7.10.2 SQL Based Knowledge Representation _____	192
7.10.3 Rule Based Engine using SQL _____	192
7.10.4 RBS in Application Domain of Medical Billing _____	193
7.10.5 Centralized Data Consistency Checks _____	193
7.10.6 Performance Evaluation Measures _____	193
7.11 Future Directions _____	194
7.11.1 Improving Time Complexity of Rule Engine _____	194
7.11.2 Improving Time Complexity of Proposed Mining Algorithm _____	194
7.11.3 Other Application Domain _____	195
7.11.4 NLP Based Learning Module _____	195
7.12 Chapter Summary _____	196
References _____	197
Appendix A: Dataset 1 _____	204