

TABLE OF CONTENTS

CHAPTER NO.	TITLE	PAGE NO.
	ABSTRACT	v
	LIST OF TABLES	xiv
	LIST OF FIGURES	xvi
	LIST OF SYMBOLS AND ABBREVIATIONS	xix
1	INTRODUCTION	1
1.1	TEMPORAL DATA IN CLINICAL DOMAIN	2
1.1.1	Temporal Data Categories	4
1.2	TEMPORAL DATA PRE-PROCESSING	5
1.2.1	Data Cleaning	6
1.2.2	Data Integration and Transformations	7
1.2.3	Temporal Data Representation	7
1.2.3.1	Time series dimensionality reduction	8
1.2.3.2	Time series segmentation	8
1.2.3.3	Temporal data abstraction	9
1.2.4	Attribute Selection	10
1.3	TEMPORAL DATA MINING	12
1.3.1	Clustering	12
1.3.2	Classification	13
1.3.3	Rule Discovery	13
1.4	RESEARCH MOTIVATION	14
1.5	RESEARCH OUTCOME	14
1.5.1	Research Contributions and Findings	14
1.5.2	Societal Contribution	15
1.6	ORGANISATION OF THESIS	15

CHAPTER NO.	TITLE	PAGE NO.
2	LITERATURE REVIEW	16
2.1	TEMPORAL REASONING IN CLINICAL DOMAIN	16
2.2	TEMPORAL DATA ABSTRACTION IN CLINICAL DATA ANALYSIS	18
2.3	TEMPORAL DATA CLASSIFICATION	19
2.4	RESEARCH DIRECTIONS	25
3	SYSTEM FRAMEWORK	29
3.1	TEMPORAL DATA PRE-PROCESSING	30
3.1.1	Missing Data Imputation	30
3.1.2	Temporal Pattern Extraction	30
3.1.3	Attribute Selection	31
3.1.4	Segmentation	31
3.2	TEMPORAL CLASSIFICATION	32
3.3	PERFORMANCE EVALUATION	34
4	DATASET DESCRIPTION	39
4.1	HEPATITIS DATASET	39
4.2	THROMBOSIS DATASET	40
4.3	PARKINSON'S DISEASE DATASET	44
5	TEMPORAL ROUGH SET INDUCED NEURO-FUZZY CLASSIFIER FOR UNEVENLY SPACED CLINICAL TIME SERIES DATA	47
5.1	SYSTEM FRAMEWORK OF TRiNF CLASSIFIER	47
5.1.1	Temporal Data Acquisition (TDA)	49

CHAPTER NO.	TITLE	PAGE NO.
	5.1.1.1 Missing data imputation and temporal pattern extraction	49
	5.1.1.2 Temporal rough set based attribute selection	55
	5.1.2 Temporal Classification using Neuro-Fuzzy	61
5.2	EXPERIMENTAL RESULTS AND DISCUSSIONS	64
5.3	CONCLUSION	73
6	Q-BACKPROPAGATED TIME DELAY NEURAL NETWORK CLASSIFIER FOR DIAGNOSING PARKINSON'S DISEASE	74
6.1	PARKINSON'S DISEASE TRAITS	74
6.2	COMPUTER ASSISTED DECISION SUPPORT SYSTEM FOR DIAGONSING PARKNISON'S DISEASE: BACKGROUND	76
6.3	METHODOLOGICAL FRAMEWORK FOR Q-BTDNN	79
6.3.1	Q-BTDNN Structure and Q-BP Learning	79
6.3.2	Application of Q-BTDNN to Diagnose Gait Disturbances in Parkinson's Disease	84
6.4	EXPERIMENTAL SETTINGS	85
6.5	RESULTS AND DISCUSSIONS	88
6.6	CONCLUSION	94
7	MISSING VALUE IMPUTATION IN UNEVENLY SPACED CLINICAL TIME SERIES DATA	95

CHAPTER NO.	TITLE	PAGE NO.
7.1	MISSING VALUE IMPUTATION	
	TECHNIQUES: BACKGROUND	95
	7.1.1 Statistical Techniques	96
	7.1.2 Machine Learning Techniques	99
7.2	TRiBS FRAMEWORK	102
	7.2.1 Significant Data Selection (SDS)	102
	7.2.2 Influence Factor Optimization and Imputation	108
7.3	EXPERIMENTAL RESULTS AND DISCUSSIONS	110
7.4	CONCLUSION	118
8	STATISTICAL TOLERANCE ROUGH SET INDUCED DECISION TREE CLASSIFIER FOR UNEVENLY SPACED CLINICAL TIME SERIES DATA	119
8.1	SYSTEM FRAMEWORK FOR STRiD CLASSIFIER	119
	8.1.1 Temporal Data Pre-Processing	120
	8.1.1.1 Missing value imputation and temporal pattern extraction using FIDES	120
	8.1.1.2 Temporal interpreted tolerance rough set attribute selection	130
	8.1.2 Temporal Classification	137
8.2	EXPERIMENTAL RESULTS AND DISCUSSIONS	139
8.3	CONCLUSION	149

CHAPTER NO.	TITLE	PAGE NO.
9	FORECAST-ERROR APPROXIMATION BASED BOTTOM-UP SEGMENTATION FOR TIME DELAY NEURAL NETWORK CLASSIFICATION	151
9.1	SYSTEM FRAMEWORK FOR FeAB SEGMENTATION AND CLASSIFICATION	151
9.1.1	Temporal Summarization	153
9.1.1.1	Temporal Smoothing: Hanzak Improved DES	153
9.1.1.2	Forecast-error approximation based bottom-up (FeAB) segmentation	157
9.1.2	Temporal Classification: Time Delay Neural Network (TDNN)	160
9.2	EXPERIMENTAL RESULTS AND DISCUSSIONS	164
9.3	CONCLUSION	170
10	CONCLUSION AND FUTURE WORKS	171
10.1	CONCLUSION	171
10.2	FUTURE WORKS	177
	REFERENCES	179
	LIST OF PUBLICATIONS	196

LIST OF TABLES

TABLE NO.	TITLE	PAGE NO.
4.1	Hepatitis and Thrombosis Dataset Summary	41
4.2	Hepatitis and Thrombosis Dataset Description	41
4.3	Parkinson's Dataset Summary	46
4.4	Parkinson's Dataset Description	46
5.1	Normalized and Discretized Temporal Patterns for Subset of Lab Examination (T-BIL, GPT)	60
5.2	Results of Wright updated DES method for Hepatitis Patient_ Mid=1, Lab Test = ALB, Year= 1981	65
5.3	Results of Temporal Attribute Selection	68
5.4	Comparison of the Classification Results for TRiNF Classifier Evaluation	69
6.1	RMSE Values Vs Epochs observed for varied combinations in Q-BTDNN Structure	89
6.2	Comparison of Classification Results for Parkinson's Disease Dataset	91
7.1	Fitness Position (Missing attribute- LAP)	112
7.2	Comparison of Imputation Results of TRiBS, KNN, EM, IDW for Hepatitis Datasets	115
7.3	Comparison of Imputation Results of TRiBS, IDW, KNN, EM, IDW for Thrombosis Datasets	116
8.1	Parameter Values for the Fuzzy Sets in FIDES	123
8.2	Temporal Pattern for Lab Test (ALB, T-CHO), Patient (1-5)	135

TABLE NO.	TITLE	PAGE NO.
8.3	Results of FIDES, Hanzak updated DES and Classical DES for Subset of Hepatitis patient (patient_ MID=1, Lab Examination = ALB, year= 1981)	141
8.4	Performance measures FIDES, Hanzak updated DES Vs Classical DES for Subset of Hepatitis and Thrombosis Dataset	142
8.5	Results of Attribute Selection using Temporal Tolerance Rough Set	146
8.6	Performance Comparison on Classification Results of STRiD, NN, C4.5 and ID3	148
8.7	Comparative Analysis based on AUC measures	149
9.1	Forecasting Results using Hanzak updated DES for Hepatitis Patients [medical identity (MID) 1, lab test-CRE, year- 2000]	166
9.2	Classification Results of TDNN using FeAB Segmentation	169

LIST OF FIGURES

FIGURE NO.	TITLE	PAGE NO.
1.1	EHR Data Sample	3
1.2	Time Series Patterns (a): Upward Trend, (b): Downward Trend, (c): No Change (d): Seasonal (e): Cycle (f): Irregular	4
1.3	Time Series Categories (a) Evenly Spaced (Regular) Time Series (b) Unevenly Spaced (Irregular) Time Series	5
1.4	Attribute Selection Categorization	11
3.1	Generic Framework of the Temporal Data Mining System for Clinical Time Series Data	29
4.1	Gait Disturbances in Parkinson's Disease	44
4.2	Wearable Sensors: (a) Wearable Sensors Attached Shoes; (b) Eight Sensors in the Shoe's Insole	45
5.1	System Framework for TRiNF Classifier	48
5.2	Illustration of Missing Data Imputation and Temporal Pattern Extraction for Hepatitis Patient (MID=1)	54
5.3	Network Training Process: (a) Network Structure and (b) Membership Function Plot for Trend Patterns of Laboratory Test	63
5.4	MSE for different Alpha and Beta: Hepatitis Patient (patient_ MID=1, Lab Examination = T-BIL)	66

FIGURE NO.	TITLE	PAGE NO.
5.5	First Level Degree of Dependency	67
6.1	Gait Cycle	75
6.2	Q-BTDNN structure	80
6.3	Total Force (a) Left Leg of PD Patient, (b) Right Leg of PD Patient (c) Left Leg of Normal Person (d) Right Leg of Normal Person	87
6.4	Q-BTDNN- Training Iterations Vs Root Mean Square Error (RMSE)	90
6.5	Results of Statistical Paired t-test on Classification Accuracies (a) Dataset: Yogev et al. 2005 (b) Dataset: Frenkel-Toledo et al. 2005b (c) Dataset: Hausdorff et al. 2007	93
7.1	TRiBS Framework	102
7.2	Illustration of Similarity Search Process in Significant Data Selection Process	107
7.3	Dependency Set for Hepatitis Attributes	111
7.4	Dependency Set for Thrombosis Attributes	112
7.5	Plot on Fitness Value Vs Iterations (Missing attribute-LAP)	113
7.6	Comparison of Classification Accuracies: (a) NN with TRiBS, KNN, EM and IDW (b) DT with TRiBS, KNN, EM and IDW and (c) SVM with TRiBS, KNN, EM and IDW	117
8.1	System Framework for STRiD Classifier	120
8.2	Block diagram for FIDES Method	122
8.3	Membership Function Plots (a) Input: 'Trend', (b) Input: 'Duration', (c) Output: 'Trend-Parameter'	124

FIGURE NO.	TITLE	PAGE NO.
8.4	Membership Function Plots (a) Input: 'Error-rate', (b) Input: 'Duration', (c) Output: 'mean - parameter'	125
8.5	Surface Plot for Trend FIS	128
8.6	Surface Plot for Mean FIS	128
8.7	Plot to Depict MSE for Different Alpha and Beta : Hepatitis Patient Record Subset (Patient_ MID=1, Lab Examination = ALB)	143
8.8	Comparison of performance measures on FIDES, DES, Hanzak and Wright updated DES (a) MSE and MAD (b) % Error and MAPE	144
8.9	Temporal First Mean Degree of Dependency for Hepatitis Dataset	146
9.1	System Framework for FeAB Segmentation and Classification	152
9.2	Illustration of Segmentation Process	158
9.3	Time Delay Neural Network Structure for classification	161
9.4	Plot for Segmentation Approximation Error Vs Number of Segments	167