

Contents

<i>Abstract</i>	<i>i</i>
<i>Acknowledgements</i>	<i>iii</i>
<i>Contents</i>	<i>v</i>
<i>List of Figures</i>	<i>vii</i>
<i>List of Tables</i>	<i>xi</i>
<i>List of Abbreviations</i>	<i>xiii</i>
Chapter 1: Introduction	1
1.1 Fundamentals of Multi-omics	1
1.1.1 <i>Different Types of Omics Data</i>	2
1.1.2 <i>Omics Profiling and Data Throughput</i>	3
1.2 Importance of Computation in Cancer Multi-omics Research	5
1.2.1 <i>Multi-omics Data Extraction and Processing</i>	5
1.2.2 <i>Multi-omics Data Management</i>	8
1.2.3 <i>Multi-omics Data Analysis</i>	8
1.3 Challenges in Multi-omics Studies	10
1.4 Thesis Statement	13
1.5 Organization and Scope of this Thesis	14
Chapter 2: Literature Review	19
2.1 Utilizing Multi-omics Data to Infer Biological Insights	19
2.1.1 <i>Identification of miRNA-mRNA Regulatory Modules</i>	19
2.1.2 <i>Cancer Subtypes Identification</i>	28
2.2 Conclusion	35
Chapter 3: Identification of Functional miRNA-mRNA Regulatory Modules	37
3.1 Introduction	37
3.2 Data Sets Used	38
3.3 Proposed miRNA-mRNA Module Identification Algorithm	38
3.3.1 <i>Relevant and Functionally Consistent miRNA-mRNA Module</i>	38
3.4 Experimental Setup	42
3.4.1 <i>Optimum Value of Threshold Delta</i>	42
3.4.2 <i>Optimum Value of MISIM Threshold</i>	42
3.4.3 <i>Comparison of Module Sizes and Connectivities</i>	44
3.4.4 <i>Evaluating Modules by Functional Enrichment</i>	46
3.4.5 <i>Prognostic Value of Modules in Cervical Cancer</i>	48
3.4.6 <i>Biological Significance Analysis of miRNA-Target Interactions</i>	50
3.5 Conclusion	52
Chapter 4: A Feature Weighing-Assisted Approach for Cancer Subtypes Identification	55
4.1 Introduction	55
4.2 Data Sets	56
4.3 Proposed Method	58
4.3.1 <i>Proposed Algorithm</i>	58
4.3.2 <i>Weighted Sample Similarity Network Generation</i>	61
4.3.3 <i>Similarity Network Fusion and Sample Clustering</i>	62
4.4 Experimental Setup	64
4.5 Results	65
4.5.1 <i>Quantitative Measures</i>	65
4.5.2 <i>Comparative Performance Analysis</i>	66
4.5.3 <i>Differential Expression Analysis</i>	68
4.6 Conclusion	72
Chapter 5: Recursive Integration of Synergised Graph Representations for Cancer Subtypes Identification	77
5.1 Introduction	77
5.2 Proposed Approach for Cancer-subtypes Identification	78
5.2.1 <i>Gramian Matrix and Kernel Trick</i>	78
5.2.2 <i>Graph Laplacian</i>	81

5.2.3	<i>RISynG Algorithm</i>	83
5.2.4	<i>Significance of Proposed Algorithm</i>	86
5.3	Description of Data-sets	87
5.4	Experimental Results and Discussion	88
5.4.1	<i>Performance Analysis on Multi-omics Cancer Data-sets</i>	89
5.5	Conclusion	101
Chapter 6: Capturing the Latent Space of an Autoencoder for Multi-omics Integration and Cancer Subtyping		109
6.1	Introduction	109
6.2	Materials and Method	110
6.2.1	<i>Benchmark Cancer Datasets</i>	110
6.2.2	<i>Proposed Framework</i>	110
6.3	Experimental Settings	115
6.4	Experimental Results	116
6.4.1	<i>Cluster Evaluation</i>	116
6.4.2	<i>Evaluation Based on Survival Pattern</i>	118
6.4.3	<i>Comparing Single-omic Versus Multi-omics Clustering</i>	119
6.4.4	<i>Biological Enrichment Analyses</i>	119
6.5	Conclusion	123
Chapter 7: Conclusion and Future Directions		129
7.1	Major Contributions	129
7.2	Future Directions	130
AnnexureA: Cluster Evaluation Indices		133
A.1	Internal Cluster Evaluation Indices	133
A.2	External Cluster Evaluation Indices	134
AnnexureB: Supplementary Information for Chapter 3		135
B.1	Details of Twenty two FMRMs identified by RFCM ³	135
AnnexureC: Supplementary Information for Chapter 4		147
C.1	Description of feature weighting approaches used in this study	147
C.1.1	<i>Entropy based feature weighting (EB)</i>	147
C.1.2	<i>Space filling concept (SFC)</i>	147
C.1.3	<i>Weighted similarity network fusion (WSNF)</i>	148
C.2	Sign test	149
AnnexureD: Supplementary Information for Chapter 6		155
D.1	Evaluation of Cancer Subtypes Based on Survival Pattern	155
D.2	Effect of other Unsupervised Feature Selection Approaches	155
D.3	Effect of Feature Selection based on COX proportional hazard (CPH) model on other Cancer Subtyping Methods	159
Bibliography		163