

Ch. Satyanarayana · Kunjam Nageswara Rao
Richard G. Bush

Computational Intelligence and Big Data Analytics

Applications in Bioinformatics

 Springer

Ch. Satyanarayana
Department of Computer Science
and Engineering
Jawaharlal Nehru Technological
University
Kakinada, Andhra Pradesh, India

Richard G. Bush
College of Information
Technology
Baker College
Flint, MI, USA

Kunjam Nageswara Rao
Department of Computer Science
and Systems Engineering
Andhra University
Visakhapatnam, Andhra Pradesh, India

ISSN 2191-530X ISSN 2191-5318 (electronic)
SpringerBriefs in Applied Sciences and Technology
ISSN 2196-8845 ISSN 2196-8853 (electronic)
SpringerBriefs in Forensic and Medical Bioinformatics
ISBN 978-981-13-0543-6 ISBN 978-981-13-0544-3 (eBook)
<https://doi.org/10.1007/978-981-13-0544-3>

Library of Congress Control Number: 2018949342

© The Author(s) 2019

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd. The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Contents

1	A Novel Level-Based DNA Security Algorithm Using DNA Codons	1
1.1	Introduction	1
1.2	Related Work	2
1.3	Proposed Algorithm	3
1.3.1	Encryption Algorithm	4
1.3.2	Decryption Algorithm	5
1.4	Algorithm Implementation	5
1.4.1	Encryption	5
1.4.2	Decryption	7
1.5	Experimental Results	8
1.5.1	Encryption Process	8
1.5.2	Decryption Process	8
1.5.3	Padding of Bits	9
1.6	Result Analysis	12
1.7	Conclusions	13
	References	13
2	Cognitive State Classifiers for Identifying Brain Activities	15
2.1	Introduction	15
2.2	Materials and Methods	16
2.2.1	fMRI-EEG Analysis	16
2.2.2	Classification Algorithms	17
2.3	Results	19
2.4	Conclusion	19
	References	19

3	Multiple DG Placement and Sizing in Radial Distribution System Using Genetic Algorithm and Particle Swarm Optimization	21
3.1	Introduction	21
3.2	DG Technologies	22
3.2.1	Number of DG Units	23
3.2.2	Types of DG Units	23
3.3	Mathematical Analysis	23
3.3.1	Types of Loads	23
3.3.2	Load Models	23
3.3.3	Multi-objective Function (MOF)	24
3.3.4	Evaluation of Performance Indices Can Be Given by the Following Equations	25
3.4	Proposed Methods	26
3.4.1	Genetic Algorithm (GA)	26
3.4.2	Particle Swarm Optimization (PSO)	26
3.5	Results and Discussions	26
3.5.1	33-Bus Radial Distribution System	26
3.5.2	69-Bus Radial Distribution System	29
3.6	Conclusions	34
	References	35
4	Neighborhood Algorithm for Product Recommendation	37
4.1	Introduction	37
4.2	Related Work	38
4.3	Existing System	39
4.4	Proposed System	41
4.5	Experiments and Results	47
4.6	Conclusion and Future Work	51
	References	52
5	A Quantitative Analysis of Histogram Equalization-Based Methods on Fundus Images for Diabetic Retinopathy Detection	55
5.1	Introduction	55
5.1.1	Extracting the Fundus Image From Its Background	56
5.1.2	Image Enhancement Using Histogram Equalization-Based Methods	57
5.2	Image Quality Measurement Tools (IQM)—Entropy	59
5.3	Results and Discussions	59
5.4	Conclusion	61
	References	62

6	Nanoinformatics: Predicting Toxicity Using Computational Modeling	65
6.1	Introduction	65
6.2	Identification of Properties	66
6.2.1	Physicochemical Properties	66
6.2.2	Theoretical Chemical Descriptor	67
6.3	Computational Techniques	69
6.4	Prediction on the Basis of Live Cells	70
6.5	Experimental Analysis	70
6.6	Affirmation of the Model	71
6.7	Conclusion	71
	References	72
7	Stock Market Prediction Based on Machine Learning Approaches	75
7.1	Introduction	75
7.2	Literature Review	76
7.3	Conclusion	78
	References	79
8	Performance Analysis of Denoising of ECG Signals in Time and Frequency Domain	81
8.1	Introduction	81
8.2	Denoising	82
8.3	Denoising Filters	83
8.4	Proposed Algorithm in Time Domain	86
8.5	Denoising in Frequency Domain	87
8.6	Proposed Algorithm in Frequency Domain	88
8.7	Results and Discussion	89
8.8	Conclusion	94
	References	94
9	Design and Implementation of Modified Sparse <i>K</i>-Means Clustering Method for Gene Selection of T2DM	97
9.1	Introduction	97
9.2	Importance of Genetic Research in Human Health	99
9.3	Dataset Description	99
9.4	Implementation of Existing <i>K</i> -Means Clustering Algorithm	100
9.5	Implementation of Proposed Modified Sparse <i>K</i> -Means Clustering Algorithm	101
9.6	Results and Discussion	102
9.6.1	Cluster Error Analysis	102

9.6.2	Selection of More Appropriate Gene from Cluster Vectors	102
9.7	Conclusion	104
	References	106
10	Identifying Driver Potential in Passenger Genes Using Chemical Properties of Mutated and Surrounding Amino Acids	107
10.1	Introduction	107
10.2	Materials and Methods	108
10.2.1	Dataset Specification	108
10.2.2	Computational Methodology	109
10.3	Results and Discussions	110
10.3.1	Mutations in Both the Driver and Passenger Genes	110
10.3.2	Block-Specific Comparison Driver Versus Passenger Protein	112
10.4	Conclusion	116
	References	117
11	Data Mining Efficiency and Scalability for Smarter Internet of Things	119
11.1	Introduction	119
11.2	Background Work and Literature Review	120
11.3	Experimental Methodology	121
11.4	Results and Analysis	121
11.4.1	Execution Time	122
11.4.2	Machine Learning Models	122
11.5	Conclusion	124
	References	124
12	FGANN: A Hybrid Approach for Medical Diagnosing	127
12.1	Introduction	127
12.2	Preprocessing	130
12.3	Genetic Algorithm-Based Feature Selection	131
12.4	Artificial Neural Network-Based Classification	132
12.5	Experimental Results and Analysis	134
12.6	Conclusion	135
	References	136