Health 5.0 and Innovative, Intelligent, Biomedical Methods



Health 5.0 and Innovative, Intelligent, Biomedical Methods

Lalit Garg,

Senior Lecturer, University of Malta, Malta

Honorary Lecturer, University of Liverpool, UK

e-mail: lalit.garg@um.edu.mt

web: http://lalitgarg.info/

Phone: +356-2340-2112

20/04/2022



Some Artifiial Intelligence Systems for COVID-19 Outbreak Management









20/04/2022

Health 5.0



Kowalkiewicz, M. (2017), Health 5.0: the emergence of digital wellness, [Online] https://medium.com/qut-cde/health-5-0-the-emergence-of-digital-wellness-b21fdff635b9



Health 5.0

Features:

- Quality of life
- Self-service
- Data-driven healthcare
- Electronic life records (Patient lifelog)
- Highly proactive
- Industry agnostic
- Value based business model
- Boundaryless

Kowalkiewicz, M. (2017), Health 5.0: the emergence of digital wellness, [Online] <u>https://medium.com/qut-cde/health-5-0-the-emergence-of-digital-wellness-b21fdff635b9</u>



Health 5.0

Technologies:

- Extended Health Intelligence: AI/ML predictive/prescriptive analytics and digital twin
- Sensors: Smart wearables, cyborg, satellites, wireless sensor networks and IoTs
- 5G/6G: Cloud computing, telemedicine and Mobile health
- Robotics (computer vision and natural language processing)
- Gene editing, genomics, epigenomics proteomics and metabolomic



Predictive and prescriptive analysis in healthcare

20/04/2022



Predictive analysis in healthcare

- Predicting health status, disease risk and response based on the information already known.
- Predicting the positive and negative impact of various factors including
 - Genetic make-up or genomic profile
 - Present health status (phenotype)
 - Environmental factors (weather, pollution, geography etc)
 - Diet and Medicine
 - Sleep, stress, life style and activities (including contact tracing)
 - Exercise and therapy
 - Relationships, social and economical status





Prescriptive analysis in healthcare

Using predictive analysis to generate (personalized) prescriptions to minimize disease risk and improve or ensure health and wellbeing. Prescriptions might include:

- Controlling environmental factors (weather, pollution, geography etc)
- Diets, Medicines
- Sleep, stress, life style changes
- Exercise and therapy
- Medical tests, investigations
- Treatment, surgery, gene editing
- Precautions and monitoring

4/20/2022



20/04/2022



The Digital Twin of Earth

- This is a project lunched by ESA in 2020.
- Implement Earth observation into the creation of a digital twin of Earth – a dynamic, digital replica of our planet which accurately mimics Earth's behavior.



What is Digital Twin of Earth (Cont ..)

• Objectives :

- Monitoring environmental threats like, rising sea levels, increasing ocean acidification and more intense extreme events like floods and heat waves.
- Accurate Modeling of the Earth.
- Study the environmental activities like Forest erosion, Hydrology, Antarctica, Food Systems, Ocean and Climate Hot Spots.
- Observed near future effects of climate change on Mother Earth.





Other Applications



Fusion of Digital Twin of Earth with Human Digital Twin



GOVERNMENT OF MALTA MINISTRY FOR RESEARCH, INNOVATION AND THE CO-ORDINATION OF POST COVID-19 STRATEGY



The Malta Council for Science & Technology





esa



The complex simulation of our planet's system



L-Università ta' Malta

Digital Twin in Healthcare

- Digital twins are digital representations of human physiology built on computer models in which data relating to both the individual and the population are introduced.
- The use of digital twins in healthcare is revolutionizing clinical processes and hospital management by enhancing medical care with digital tracking and advancing modeling of the human body.
- In the future may also help physicians optimize the performance of patient-specific treatment plans.



Medical Devices Attributes For Digital Twin



DIGITAL TWIN



Facility Attributes For Digital Twin







Source: https://www.verywellhealth.com/digital-twincomputer-model-of-patients-5120469



Digital Twin in Healthcare



Human Attributes For Digital Twin





Working Procedure





Digital Twin Model





L-Università ta' Malta

Making and keeping your digital copy (software replicas) containing all required information about you together with your 3D image/scan. It can also contain:

- 1. Your general information (hight, weight, shape, BMI etc.)
- 2. Your genomic profile and phenotype
- 3. Your medical history
- 4. Your family history (health related)
- 5. Your diet, activity, travel history
- 6. Your predictive/prescriptive analysis

You can keep it and share it with your doctors or others by selecting the information you want to share.



It can help:

- 1. Predicting health status
- 2. Predicting disease risks
- 3. Recognising unexpected health changes
- 4. Identifying factors affecting health
- 5. Determining the most effecting
 - 1. Treatment
 - 2. Therapy
 - 3. Behavioural intervention (behavioural medicine)
 - 4. Dietary intervention (nutrition)





It can help:

- 6. Improving patient monitoring
- 7. Disease warning
- 8. Patient education
- 9. Improving physician-patient relationship
- 10. Predicting resource requirements
- 11. Ensuring resource availability
- 12. Focusing on Prevention
- 13. Reducing hospital visits
- 14. Facilitating effecting telemedicine

4/20/2022



It can help:

- 15. Medicine as an Open Loop System
- 16. Better personalized medicine
- 17. Reducing medical errors
- 18. Controlling drug resistance
- 19. Controlling hospital acquired infection
- 20. Reducing insurance premium
- 21. Developing and sharing patient lifelog
- 22. Promptly recognising and containing infectious disease outbreak



4/20/2022

It can help:

- 23. Promptly recognising factors responsible for noncommunicable diseases
- 24. Developing better public health measures/policies
- 25. Analysing impact of public health measures
- 26. Health fraud detection and prevention.





Received August 26, 2020, accepted August 28, 2020, date of publication August 31, 2020, date of current version September 14, 2020. Digital Object Identifier 10.1109/ACCESS.2020.3020513

Anonymity Preserving IoT-Based COVID-19 and **Other Infectious Disease Contact Tracing Model**

LALIT GARG^[0], (Member, IEEE), EMEKA CHUKWU^[0], (Graduate Student Member, IEEE), NIDAL NASSER^{©2}, (Senior Member, IEEE), CHINMAY CHAKRABORTY^{©3},

AND GAURAV GARG¹⁰⁴, (Member, IEEE)

1 Department of Computer Information System (CIS), Faculty of Information Communication Technology (ICT), University of Malta, 2080 Msida, Malta ²College of Engineering, Alfaisal University, Riyadh 50927, Saudi Arabia ³Department of Electronics and Communication Engineering, Birla Institute of Technology at Mesra, Ranchi 835215, India ⁴ABV-Indian Institute of Information Technology and Management, Gwalior 474015, India

Corresponding author: Lalit Garg (lalit.garg@um.edu.mt)

This work was supposed in part by Alfaisal University under Internal Research Grant C20220 and in part by the University of Malta's Research Innovation & Development Trust (RIDT) under Research Grant E18LO77-01.

ABSTRACT Automated digital contact tracing is effective and efficient, and one of the non-pharmaceutical complementary approaches to mitigate and manage epidemics like Coronavirus disease 2019 (COVID-19). Despite the advantages of digital contact tracing, it is not widely used in the western world, including the US and Europe, due to strict privacy regulations and patient rights. We categorized the current approaches for contact tracing, namely: mobile service-provider-application, mobile network operators' call detail, citizenapplication, and IoT-based. Current measures for infection control and tracing do not include animals and moving objects like cars despite evidence that these moving objects can be infection carriers. In this article, we designed and presented a novel privacy anonymous IoT model. We presented an RFID proof-of-concept for this model. Our model leverages blockchain's trust-oriented decentralization for on-chain data logging and retrieval. Our model solution will allow moving objects to receive or send notifications when they are close to a flagged, probable, or confirmed diseased case, or flagged place or object. We implemented and

4/20/2022



Funding body: Malta's Research Innovation & Development Trust (RIDT) and Alfaisal University, Riyadh, Saudi Arabia





Collaborative partners:

Lalit Garg¹, Emeka Chukwu¹, Nidal Nasser², Chinmay Chakraborty³, Gaurav Garg⁴

¹University of Malta

²Alfaisal University, Riyadh, Saudi Arabia

³Birla Institute of Technology, Ranchi, India

⁴ABV-Indian Institute of Information Technology and Management, Gwalior, India



L-Università ta' Malta

- 1. Preserves users privacy/anonymity
- 2. Trace any moving objects including
 - I. Human
 - II. Animals
 - III. Vehicles
- 3. Based on RFID (Radio Frequency Identification Device)
- 4. leverages blockchain's trust-oriented decentralization for on-chain data logging and retrieval



- 1. Leverages blockchain's trust-oriented decentralization for on-chain data logging and retrieval
- 2. Allows moving objects to receive or send notifications when they are close to
 - I. a flagged, probable, or
 - II. confirmed diseased case, or
 - III. flagged place or object.



- Less than one-second deployment and call time for smart contracts
- 2. 25 seconds on Ethereum public blockchain
- 3. Easy to identify clusters of infection contacts
- 4. Help deliver a notification for mass isolation while preserving individual privacy


Anonymity Preserving IoT-Based Contact Tracing Model

- 1. It can be used to
 - 1. Understand better human connectivity,
 - 2. Model similar other infection spread network,
 - 3. Develop public policies to control the spread of COVID-19
 - 4. Finding super spreader, hotspots and responsible behaviour.



Anonymity Preserving IoT-Based Contact Tracing Model





Anonymity Preserving IoT-Based Contact Tracing Model





Covid-19: Disease Network

- 1. Design
- 2. Monitoring
- 3. Tracking
- 4. Sharing
- 5. Analysing
- 6. Learning
- 7. Predicting
- 8. Preparing



Covid-19: Patient lifelog sharing

- 1. Medical history
- 2. Travel history
- 3. Activities and behaviour
- 4. Family history



More info...

• Garg L, Chukwu E, Nidal N, Chakraborty C, Garg G (2020). Anonymity preserving IoT-based COVID-19 and other infectious disease contact tracing model. IEEE Access.



Healthcare self-service Kiosk

Lalit Garg, Emeka Chukwu





Healthcare self-service Kiosk

- 1. A novel healthcare self-service Kiosk.
- 2. The kiosk is interfaced to a database, backend, and Frontend interfaces.
- 3. The kiosk is used by a client visiting the hospital to self-centre historical data.
- 4. The sensors help capture the blood pressure, temperature, and weight in this initial prototype.
- 5. Our prototype use a Raspberry pi lightweight server, connected to all the interfaces



Healthcare self-service Kiosk



Hospital bed occupancy and requirements forecasting



Intelligent Patient Management and Resource Planning for Complex, Heterogeneous, and Stochastic Healthcare Systems

Lalit Garg, Member; IEEE, Sally I. McClean, Maria Barton, Brian J. Meenan, and Ken Fullerton

Abstract-Effective resource requirement forecasting is necessary to reduce the escalating cost of care by ensuring optimum utilization and availability of scarce health resources. Patient hospital length of stay (LOS) and thus resource requirements depend on many factors including covariates representing patient characteristics such as age, gender, and diagnosis. We therefore propose the use of such covariates for better hospital capacity planning. Likewise, estimation of the patient's expected destination after discharge will help in allocating scarce community resources. Also, probable discharge destination may well affect a patient's LOS in hospital. For instance, it might be required to delay the discharge of a patient so as to make appropriate care provision in the community. A number of deterministic models such as ratio-based methods have failed to address inherent variability in complex health processes. To address such complexity, various stochastic models have therefore been proposed. However, such models fail to consider inherent heterogeneity in patient behavior. Therefore, we here use a phase-type survival tree for groups of patients that are homogeneous with respect to LOS distribution, on the basis of covariates such as time of admission, gender, and disease diagnosed; these homogeneous groups of patients can then model patient flow through a care system following stochastic pathways that are characterized by the covariates. Our phase-type model is then extended by further growing the survival tree based on covariates

provide a stochastic approach to capacity planning across complex heterogeneous care systems. The approach is illustrated using a five year retrospective data of patients admitted to the stroke unit of the Belfast City Hospital.

Index Terms—Capacity planning, cost, decision-making, forecasting, health information management, medical information systems, operations research, optimal control, prognostics and health management, stochastic systems.

I. INTRODUCTION

E FFECTIVE resource requirement forecasting is necessary to minimize the escalating cost of care by ensuring optimum utilization and availability of scarce health resources [1]. Patient hospital length of stay (LOS) and thus resource requirements depend on many factors including covariates representing patient characteristics such as age, gender, and diagnosis [2]. It is therefore necessary to consider the effect of such covariates for better capacity planning. Information about the patient demography helps in making better allocation of scarce resources. Predicting different treatment outcome,

A non-homogeneous discrete time Markov model for admission scheduling and resource planning in a cost or capacity constrained healthcare system

Lalit Garg · Sally McClean · Brian Meenan · Peter Millard

Received: 5 March 2009 / Accepted: 23 October 2009 © Springer Science+Business Media, LLC 2009

Abstract Healthcare resource planners need to develop policies that ensure optimal allocation of scarce healthcare resources. This goal can be achieved by forecasting daily resource requirements for a given admission policy. If resources are limited, admission should be scheduled according to the resource availability. Such resource availability or demand can change with time. We here model patient flow through the care system as a discrete time Markov chain. In order to have a more realistic representation, a nonhomogeneous model is developed which incorporates timedependent covariates, namely a patient's present age and the present calendar user. The model presented in this paper can using a historical dataset from the geriatric department of a London hospital.

Keywords Resource management · Admission scheduling · Non-homogeneous Markov model · Stochastic optimal control

1 Introduction

Admission scheduling [1, 2] and resource planning [3] are fundamental problems which require complex strategies to effectively manage care services ensuring entimum utilize INFORMATICA, 2011, Vol. 22, No. 1, 57–72 © 2011 Vilnius University

Phase-Type Survival Trees and Mixed Distribution Survival Trees for Clustering Patients' Hospital Length of Stay

Lalit GARG¹, Sally McCLEAN¹, Brian J. MEENAN¹, Peter MILLARD²

¹ University of Ulster, Coleraine

Co. Londonderry, BT52 1SA, UK

² St. George's Hospital Medical School

12 Cornwall Road, Cheam, Sutton, Surrey, SM2 6DR, UK

e-mail: (l.garg, si.mcclean, bj.meenan)@ulster.ac.uk, phmillard@tiscali.co.uk

Received: October 2009; accepted: January 2011

Abstract. Clinical investigators, health professionals and managers are often interested in developing criteria for clustering patients into clinically meaningful groups according to their expected length of stay. In this paper, we propose two novel types of survival trees; phase-type survival trees and mixed distribution survival trees, which extend previous work on exponential survival trees. The trees are used to cluster the patients with respect to length of stay where partitioning is based on covariates such as conder, are at the time of admission and primary diagnosis code. Likelihood

57

IMA Journal of Management Mathematics (2009) 20, 327–344 doi:10.1093/imaman/dpn030 Advance Access publication on November 20, 2008

Non-homogeneous Markov models for sequential pattern mining of healthcare data

LALIT GARG[†] AND SALLY MCCLEAN[‡]

School of Computing and Information Engineering, University of Ulster, Coleraine, Co. Londonderry, BT52 ISA, UK

BRIAN MEENAN§

School of Engineering, University of Ulster, Jordanstown Campus, Newtownabbey, Co. Antrim, BT37 0QB, UK

AND

PETER MILLARD

St. George's Hospital Medical School, 12 Cornwall Road, Cheam, Sutton, Surrey SM2 6DR, UK

[Received May 2007; accepted May 2008]

Sequential pattern mining has been a popular data mining technique for extracting useful information from large databases and has successfully been used for numerous industrial and commercial problems. This paper presents a new mathematical modelling application to healthcare, providing important information to health service managers and policy makers to help them identify sequential patterns which require attention for efficiently managing scarce healthcare resources and developing effective healthcare management policies. In healthcare, these sequential patterns are analogous to the patient pathways. We present a non-homogeneous Markov model for identifying not only patient pathways which have high probability but also for identifying pathways which incur high cost or time. In order to have a more

Hospital bed occupancy and requirements forecasting

•**Collaborative partners:** Nanyang Technological University and Tan Tock Seng Hospital, Singapore.

•Approach: Markov modelling, reinforcement learning

•Data: Tan Tock Seng Hospital, Singapore.



Hospital bed requirements forecasting using satellite, weather & air quality data





- Garg L, McClean SI, Meenan BJ, Barton M, Fullerton K (2012). Intelligent patient management and resource planning for complex, heterogeneous, and stochastic healthcare systems. In press. IEEE Transactions on Systems, Man, and Cybernetics--Part A: Systems and Humans 42(6): 1332 1345. Impact Factor: 2.093, Journal H-Index(SJR:1996-2011): 54, SJR(2011):0.085, Scopus SNIP(2011):3.594, Scopus citations: 2, ISI Web of Knowledge Citations: 1, scHolar index: 2.
- Garg L, McClean SI, Meenan BJ, Barton M, Fullerton K (2011). An Extended Mixture Distribution Survival Tree for Patient Pathway Prognostication. Communications in Statistics: Theory and Methodology. 42(16):2912-2934. Impact Factor: 0.351, Journal H-Index(SJR:1996-2011): 24, SJR(2011):0.035, Scopus SNIP(2011):0.550.
- Barton M, McClean SI, Gillespie J, Garg L, Wilson D, Fullerton K (2012). Is it beneficial to increase the provision of thrombolysis?- A discrete-event simulation model, QJM: An International Journal of Medicine.105(7),(art. no. hcs036):665-673. Impact Factor: 2.146, Journal H-Index(SJR:1993-2011): 71, SJR(2011):0.18, Scopus SNIP(2011):0.897, Scopus citations: 3, ISI Web of Knowledge Citations: 1,scHolar index: 2.
- Garg L, McClean SI, Meenan BJ, Millard PH (2011). Phase-type survival trees and mixed distribution survival trees for clustering patients' hospital length of stay. INFORMATICA. 22(1): 57-72. Impact Factor: 1.786, Journal H-Index(SJR:1996-2010): 15, SJR(2011):0.048, Scopus SNIP(2011):1.814, Scopus citations: 8, ISI Web of Knowledge Citations: 7, scHolar index: 10.



- McClean SI, Barton M, Garg L, Fullerton K (2011). A Modeling framework that combine Markov models and discrete event simulation for stroke patient care. ACM Transactions on Modelling and Computer Simulation (TOMACS), 21(4). Impact Factor: 0.684, Journal H-Index(SJR:1991-2009): 26, SJR(2011): 0.043, Scopus SNIP(2011):2.053, ISI Web of Knowledge Citations: 2, Scopus citations: 4, scHolar citations: 6.
- Jahangirian M, Eldabi T, Garg L, Jun T, Naseer A, Patel B, Stergioulas L, Young TP (2011). A Rapid Review Method For Extremely Large Corpora of Literature: Applications to the domains of Modelling, Simulation, and Management. International Journal of Information Management (IJIM), 31(3):234-243. Impact Factor: 1.554, Journal H-Index(SJR:1986-2010): 32,SJR(2011):0.043, Scopus SNIP(2011):2.517, Scopus citations: 3,ISI Web of Knowledge Citations: 2, scHolar citations: 5.
- Gillespie J, McClean SI Scotney B, Garg L, Barton M, Fullerton K (2011). Costing Hospital Resources for Stroke Patients using Phase-type Models. Health Care Management Science. 14(3):279-291. Impact Factor: 0.394, SJI(2011): 0.063, Scopus SNIP(2011): 1.169, Journal H-Index(SJR:1998-2009): 23, Scopus citations: 1, scHolar citations: 3.
- Papadapoulou A, Tsaklidis G, McClean SI, Garg L (2011). On the moments and the distribution of the cost of a semi Markov model for healthcare systems. Methodology and Computing in Applied Probability.14(3):717-737. Impact Factor: 0.774, Journal H-Index(SJR:2004-2010): 11, SJR(2011):0.040, Scopus SNIP(2011):0.718.



- Garg L, McClean SI, Meenan BJ, Millard PH (2010). A non-homogeneous discrete time Markov model for admission scheduling and resource planning in a care system. Health Care Management Science. 13(2):155–169. Impact Factor: 0.394, Journal H-Index(SJR:1998-2009): 23, SJI(2011):0.063, Scopus SNIP(2011):1.169, Scopus citations: 8, ISI Web of Knowledge Citations: 5, scHolar citations: 18.
- Garg L., McClean SI, Barton M, Meenan BJ and Fullerton K (2010). The extended mixture distribution survival tree based analysis for clustering and patient pathway prognostication in a stroke care unit. International Journal of Information Sciences and Application. 2(4): 671-675.
- Garg L, McClean SI, Meenan BJ, Millard PH (2009). Non-homogeneous Markov Models for Sequential Pattern Mining of Healthcare Data. IMA journal Management Mathematics. 20(4): 327-344. Journal H-Index(SJR:2001-2009): 11,SJR(2011):0.033, Scopus SNIP(2011):0.671, Scopus citations: 1, ISI Web of Knowledge Citations: 2, scHolar citations: 4.
- Garg L, McClean SI, Barton M (2008). Is Management Science doing Enough to Improve Healthcare?. International Journal of Human and Social Sciences 3(2): 99-103. scHolar citations: 4.
- Garg L, McClean SI, Meenan BJ, Millard PH (2009). A phase type survival tree model for clustering patients according to their hospital length of stay. Proceedings of the XIII International Conference on Applied Stochastic Models and Data Analysis (ASMDA 2009), pp. 497-502, Eds: Leonidas Sakalauskas, Christos Skiadas, Edmundas K. Zavadskas, ISBN: 978-9955-28-463-5, Publisher: Vilnius Gediminas Technical University Press. scHolar citations: 3.



- Barton M, McClean SI, Garg L, Fullerton K (2009). Modelling Stroke Patient Pathways using Survival Analysis and Simulation Modelling. Proceedings of the XIII International Conference on Applied Stochastic Models and Data Analysis (ASMDA 2009), pp. 370-373, Eds: Leonidas Sakalauskas, Christos Skiadas, Edmundas K. Zavadskas, ISBN: 978-9955-28-463-5, Publisher: Vilnius Gediminas Technical University Press. scHolarcitations: 4.
- McClean SI, Garg L, Barton M, Fullerton K, Millard PH (2009). Using Markov Systems to plan Stroke Services. In Intelligent Patient Management (Eds. McClean, S.; Millard, P.; El-Darzi, E.; Nugent, C.D.), Book Series: Studies in Computational Intelligence, Vol. 189. pp. 241-256. Journal H-Index(SJR:1998-2010): 10, SJI(2011):0.029, Scopus SNIP(2011):0.310, Scopus citations: 2, ISI Web of Knowledge Citations: 4, scHolar citations: 3.
- McClean SI, Millard PH, Garg L (2008). Using Markov Models for Decision Support in Management of High Occupancy Hospital Care. In Intelligent Techniques and Tools for Novel System Architectures (P. Chountas, I. Petrounias, J. Kacprzyk Eds), Springer's book series Studies in Computational Intelligence (SCI). Vol. 109/2008. pp. 187–199. Journal H-Index(SJR:1998-2010): 10, SJI(2011):0.029, Scopus SNIP(2011):0.310.
- McClean SI, Garg L, Meenan BJ, Millard PH. (2007). Non-Homogeneous Markov Models for Performance Monitoring in Healthcare. (In C.H. Skiadas, Eds.) Recent Advances In Stochastic Modelling and Data Analysis. World Scientific Publishing, pp. 146-153. scHolar citations: 4.



Dr Lalit Garg Samira Cachia Spiteri Pranali Ozarkar



- Create an intelligent mobile app with the aims of:
 - Empowering patients to taking a pro-active approach with regards to their health
 - Personalising patient care
 - A more accurate patient health status by the use of remote monitoring
 - Strengthening the relationship between physicians and their patients
 - Reducing the strain on the healthcare system and its resources
- The purpose of this solution is to combine the existing technological services already offered along with machine learning techniques to make healthcare services more accessible.



- Patients Features
- Patient Dashboard
- Medication Tracker
- Appointments
- Symptom Diagnostics
- Video and Chat appointments with GP's
- Doctor Directory
- Patient Education
- Medical and Surgical History
- Disease Prevalence Statistics



- Doctors Features
- Medical records of current patients
- Medication prescription
- Tickets of referrals
- Video and Chat appointments with patients
- Patient diagnostics reports
- Alerts for irregular or life-threatening vitals
- Test scheduling



Missing data handling





Tensor-Based Methods for Handling Missing Data in Quality-of-Life Questionnaires

Lalit Garg, Member, IEEE, Justin Dauwels, Senior Member, IEEE, Arul Earnest, and Khai Pang Leong

Abstract—A common problem with self-report quality-of-life questionnaires is missing data. Despite enormous care and effort to prevent it, some level of missing data is common and unavoidable. Missing data can have a detrimental impact on the data analysis. In this paper, a novel approach to imputing missing data in quality-of-life questionnaires is proposed, based on matrix and tensor decompositions. In order to illustrate and assess those methods, two datasets are considered: The first dataset contains the responses of 100 patients to a systemic lupus erythematosus-specific quality-of-life questionnaire; the other contains the responses of 43 patients to a rhino-conjunctivitis quality-of-life questionnaire. The two datasets contain almost no missing data, and for testing purposes, data entries are removed at random to have missing completely at random data. Several proportions of missing values are considered, and for each, the imputation error is assessed through k-fold cross validation. We also evaluate different imputation methods for missing at random and missing not at randomdata. The numerical results demonstrate that the proposed tensor factorization-based methods outperform standard methods in terms of root mean square error with at least 4% improvement, while the bias and variance are similar.

Index Terms—Health information management, medical information systems, missing data imputation, quality-of-life questionnaires, tensor decomposition. treatment options for the patient [1]. However, a common problem with such questionnaires is missing data [1].

The best possible method to deal with missing data is to avoid the problem with careful planning and data collection [1], [2]. However, despite enormous care and effort to prevent it, some level of missing data is common and unavoidable [1]–[3]. Such missing data can have a detrimental impact on statistical analysis based on the questionnaires responses, including biased parameter estimates and inflated standard errors [1], [2]. Moreover, most standard data analysis techniques are developed for complete data, and they cannot directly be used with missing data [1]. A variety of methods have been suggested for imputing missing values in data [1], [4]. However, most of these methods fail to fully exploit correlations in the data, and lead to unreliable imputation of the missing values [5]. More research is desperately needed to assess and improve the reliability of missing data handling methods [1].

In this paper, we propose novel approaches for handling missing data more effectively, specifically, matrix and tensor decomposition methods. We assess and illustrate these techniques by means of two datasets. The first contains the responses of 100 patients to a systemic lupus erythematosus-specific quality-

Web-based tools for Missing data handling in medical questionnaires

Funding body: Nanyang Institute of Technology in Health & Medicine (NITHM), Singapore, University of Malta, Malta

Medical questionnaires with missing data



CP based collaborative filtering for missing data imputation









Web-based tools for Missing data handling in medical questionnaires

Collaborative partners:

Lalit Garg, Justin Dauwels¹, Arul Earnest^{2,3}, Leong Khai Pang³





²Duke-NUS Graduate Medical School, Singapore

Tan Tock Seng

³Tan Tock Seng Hospital (TTSH), Singapore





CP based missing data imputation method

						T	hird fo	ollo	ow-on									7
			SN		Q1	Q	2	Q	3 (Q 4	C	25	Q6		Q7	Q	3	08
			Patient	1	Nan		4		4		4	4		4	Na	n	0	4
			Patient	2	6	;	Nan		5		7	Nan		7		6	6	4
	ſ			Secon	d follo	w-o	on									7	2	1
	SN		Q1	Q2	Q3	(Q4	C	25	Q6		Q7	Q8		08		4	1
	Patient 1		2	2	2	2	2	2	Nan		2	Nan	_		2	2	2	3
	Patient	2	4	6	5 Na	n	e	5	2		6	5	5		5	2	5	4
		First fo	ollow-o	n	_							_			2	1	3	2
SN	Q1	Q2	Q3	Q4	Q5		Q6		Q7	Q	В	Q9			4	1	3	4
Patient 1	Nan	1		1	1	1		1		1	1		1		2	3	3	3
Patient 2	3	5	;	1	5	2		6		5	Nar	1	3		5	4	2	1
Patient 3	1	1	. :	1	1 N	an		2		2	2	2	1		3	2	2	2
Patient 4	1	Nan		1	1	1		1		1	4	Ļ	1		3	4	2	2
Patient 5	2	1	. :	1	1	1		3		2	2	2	3		3	3	1	2
Patient 6	3	4	1 1	2	5	4		7		4	5	5	4		2	1	1	4
Patient 7	1	1	. Nar	۱	1	1		1	Na	n	3	3	2		3	3	3	3
Patient 8	7	7	'	3	7	7		7		3	3	3	4		3	3	5	
Patient 9	1	1	. :	1	1	1		1		2	3	3	3		2	2		
Patient 10	6	7	'	5	7	6		7		1	2	2	1		4	4		
Patient 11	1	1	. :	1	1	1		1		3	3	3	3		1	1		
Patient 12	Nan	1	. :	1	1	1		3		3	3	3	3		5	3		
Patient 13	1	4 No.		1	3	1		3		2	2		2		/			
Patient 14	4	Nan		1	4	1		4		3	Nar	1	4		/			
Patient 15	1	1		1 Na	an	1		5		1	1		1	/				
Patient 16	1	3	3	1	1 Na	an		3		3	3	3	3					

L-Università

ta' Malta



Dr Lalit Garg

CP based missing data imputation method

		Λ	Third follow-on																
			SN		Q1		Q	2	Q3		Q4	0	25	Q6		Q7	Q	8	08
			Patien	t 1		4		4		4		4	4	1	4		3	9	4
			Patien	t 2		6		7		5		7	3	3	7		6	6	4
				Seco	ond f	ollov	v-c	on									\overline{A}	2	. 1
	SN		Q1	Q2	Q	13	C	24	Q	5	Q6		Q7	Q8		08		4	. 1
	Patient 1		2		2	2	2	2	2	2		2		2	7		2	2	. 3
	Patient	: 2	4	Ļ	6	2	2	e	5	2	2	6		5	5	5	2	5	4
		First	follow-	on											2	2	1	3	2
SN	Q1	Q2	Q3	Q4		Q5		Q6	(קב	Q	В	Q9		4	Ļ	1	3	4
Patient 1	1		1	1	1		1		1		1	1	L	1	2	2	3	2	1
Patient 2	3		5	1	5		2		6		5	3	3	3	5	>	4	2	
Patient 3	1		1	1	1		1		2		2	2	2	1	3	5	2	3	3
Patient 4	1		1	1	1		1		1		1	4	1	1	3	5	4	2	2
Patient 5	2		1	1	1		1		3		2	2	2	3	3	5	3	4	4
Patient 6	3		4	2	5		4		7		4	5	5	4	4	-	1	1	1
Patient 7	1		1	1	1		1		1		3	3	3	2	3	5	3	3	3
Patient 8	7		7	3	7		7		7		3	3	3	4	3	5	3		
Patient 9	1		1	1	1		1		1		2	3	3	3	4	-	2		/
Patient 10	6		7	6	7		6		7		1	2	2	1	4	ł	4		
Patient 11	1		1	1	1		1		1		3	3	3	3	1	-	1		
Patient 12	1		1	1	1		1		3		3	3	3	3	3)	3		
Patient 13	1		4	1	3		1		3		2	2	2	2		/			
Patient 14	4		4	1	4		1		4		3	4	1	4					
Patient 15	1		1	1	1		1		5		1	1	L	1	/				
Patient 16	1		3	1	1		1		3		3	3	3	3					

L-Università

ta' Malta



Dr Lalit Garg

More info...

Garg L, Dauwels J, Earnest A, Pang L (2013) Tensor based methods for handling missing data in quality-of-life questionnaires. IEEE Journal of Biomedical and Health Informatics. In press. doi: 10.1109/JBHI.2013.2288803. URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6656914&isnumb

er=6363502

Asif MT, Srinivasan K, Garg L, Dauwels J, Jaillet P (2013) Low-dimensional Models for Missing Data Imputation in Road Networks, ICASSP 2013, accepted. http://web.mit.edu/jaillet/www/general/missingdata_final.pdf.

http://lalitgarg.weebly.com/missingdatahandlingproject.html



More info...

Dauwels J, Garg L, Earnest A, Pang LK (2012). Tensor Factorizations for Missing Data Imputation in Medical Questionnaires, The 37th International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Kyoto, Japan, March 25 - 30, 2012.

Dauwels J, Garg L, Earnest A, Pang LK (2011). Handling Missing Data in Medical Questionnaires Using Tensor Decompositions. The Eighth International Conference on Information, Communications, and Signal Processing (ICICS 2011). Singapore 13-16 December, 2011.

http://lalitgarg.weebly.com/missingdatahandlingproject.html



MDSS for managing acute upper gastrointestinal bleeding





MDSS for managing acute upper gastrointestinal bleeding

Collaborative partners: Nanyang Technological University and Tan Tock Seng Hospital, Singapore.

Data: Tan Tock Seng Hospital, Singapore.

Approach: Pattern analysis and matching, Machine learning, rule based systems.



HIV-disease progression modelling





Using phase type distributions for modelling HIV disease progression

Lalit Garg¹, Giovanni Masala², Sally I. McClean¹, Marco Micocci², Giuseppina Cannas² ¹School of Computing and Information Engineering, University of Ulster, Coleraine, UK ²Faculty of Economics, University of Cagliari, Cagliari, Italy garg-l@email.ulster.ac.uk, si.mcclean@ulster.ac.uk, gb.masala@unica.it

Abstract

Disease progression models are useful tools for gaining a systems' understanding of the transitions to disease states, and characterizing the relationship between disease progress and factors affecting it such as patients' profile, treatment and the HIV diagnosis stage. Patients are classified into four states (based on CD4+ T-lymphocyte count) and all the transitions are allowed, Examinations to identify disease progression of the patient are carried out routinely throughout the follow-up period. Therefore, the times spent at the various HIV infection stages are interval censored or right censored. This makes difficult to use simple statistical methods such as regression to model the disease progression and its relationship with the diagnosis stage. We present a novel, more intuitive and realistic approach based on phase type distributions to

Organization (WHO) proposed a simplified model which classifies HIV infection (or the progression of HIV disease) as a four stage bidirectional process [2].

The immunological status of an HIV infected patient can not only progress sequentially from the first stage (i.e., stage 1) to the final stage (i.e., stage 4) but also regress or jump from a stage to any other stage. Examinations to identify disease progression and to determine the CD4 count of the patient are carried out routinely throughout the follow-up period. Therefore the times spent at the various HIV infection stages (HI stages) are interval censored [3] or right censored. This makes difficult to use simple statistical methods such as regression to model the disease progression and its relationship with the diagnosis stage. Markov models [4], hidden Markov models [5], and semi-Markov models [6] have been popular choices to model HIV disease progression.
HIV-disease progression modelling

Collaborative partners: University of Ulster, UK and University of Cagliari, Italy.

Approach: Phase type survival tree analysis, survival analysis, Markov process model, Bayesian Analysis

Data: Istituto Superiore di Sanità, Roma, Italy



HIV-disease progression modelling

More info:

Garg, L., Masala G., McClean S.I., Micocci M., Cannas G. (2012). Using phase type distributions for modelling HIV disease progression, Computer-Based Medical Systems (CBMS), 2012 25th International Symposium on, 20-22 June 2012. doi: 10.1109/CBMS.2012.6266408.

Garg L, McClean SI, Meenan BJ, Millard PH (2011). Phase-type survival trees and mixed distribution survival trees for clustering patients' hospital length of stay. INFORMATICA. 22(1): 57-72.



Other Innovative, Intelligent, Biomedical Methods Projects



Smart Sensor for EEG Acquisition and Epileptic Seizure Detection and prediction





Effective Data Acquisition for Machine Learning Algorithm in EEG Signal Processing

James Bonello, Lalit Garg, Gaurav Garg and Eliazar Elisha Audu

Abstract The aim of this paper is to demonstrate that small dataset can be used in machine learning for seizure monitoring and detection using smart organization of multichannel EEG sensor data. This reduces training time and improves computational performance in terms of space and time complexities on hardware implementations. The proposed approach has been tested and validated using CHB-MIT dataset containing EEG recordings of 24 clinically verified seizure and non-seizure pediatric patients. The predictability is discussed in terms of the latency and the required length of data for the proposed approach over the state-of-the-art method in the field of EEG-based seizure prediction.

Keywords EEG · Multichannel data · Sensor data · Machine learning Automated seizure detection

1 Introduction

The traditional process of manually deciphering information and analysis of electroencephalography (EEG) data for medical diagnosis is practically challenging and technically demanding to experts. With the emergence of machine learning and their applications in classification, nonlinear approximation and pattern recognition,

J. Bonello () · L. Garg University of Malta, Maida, Malta e-mail: jamesbonello9@gmail.com

L. Garg e-mail: lalit.garg@um.edu.mt

L. Garg - E.E. Audu University of Liverpool, Liverpool, UK e-mail: eliazar.audu@online.liverpool.ac.uk

G. Garg University of Ulster, Magee Campus, Londonderry, UK e-mail: garg-g@email.ulster.ac.uk

© Springer Nature Singapore Pte Ltd. 2018 M. Part et al. (eds.), Soft Computing: Theories and Applications, Advances in Intelligent Systems and Computing 584, https://doi.org/10.1007/978-981-10-5699-4_23

Smart Sensor for EEG Acquisition and Epileptic Seizure Detection

Collaborative partners: Nanyang Technological University, Singapore and Massachusetts General Hospital, MIT, USA.

Approach: Singular Vector Machine, Extreme learning machine, probabilistic clustering, Empirical mode decomposition.

Funding Body: MNN-RIDT

Data: Massachusetts General Hospital, MIT, USA.



Smart Sensor for EEG Acquisition and Epileptic Seizure Detection

More info:

Ali H. Shoeb, John V. Guttag: Application of Machine Learning To Epileptic Seizure Detection. ICML 2010: 975-982.



Electroencephalogram (EEG) Pros: Fast Temporal Response Cons: Poor Spatial Resolution (CPP and 2-D)



3D fMRI Image of Brain functional Magnetic Resonance Imaging(fMRI) Pros: Good Spatial Resolution (3D) Cons: Slow BOLD transient response

L-Università

Malta

Dr Lalit Garg







Collaborative partners: Intelligent Systems Research Centre, University of Ulster, UK, Nanyang Technological University, Singapore

Funding: Northern Ireland Department for Education and Learning

Approach: Probabilistic clustering, cluster analysis, functional analysis, convolution, SVM, ELM, factor analysis, latent class model (LCM)

L-Università

' Malta

Dr Lalit Garg



More info:

Garg G, Prasad G, **Garg L,** Coyle D (2011). <u>Gaussian Mixture</u> <u>Models for Brain Activation Detection from fMRI Data</u>, <u>International</u> <u>Journal of Bioelectromagnetism</u>. 13(4):255-260.

Garg G, Girijesh P, Damien C (2013). <u>Gaussian Mixture Model-based noise reduction in resting state fMRI data</u>. Journal of neuroscience methods. 215(1):71-77.

L-Università

Dr Lalit Garg

ta' Malta



Predicting Neurological Disorder via Social Media





L-Università ta' Malta Dr Lalit Garg

Predicting Neurological Disorder via Social Media

Collaborative partners: Jiwaji University Gwalior



Approach: CES-D screening test, Social media analytics, Major Depressive Disorder (MDD) classifier, Probabilistic clustering, cluster analysis, functional analysis, convolution, SVM, ELM, factor analysis, latent class model (LCM)

L-Università

ta' Malta

Dr Lalit Garg



A masked-image recognition system



ta' Malta

A Masked-image recognition system

Lalit Garg, Emeka Chukwu



ta' Malta





A Masked-image recognition system

- Developing novel approaches to anonymously recognizing face masked persons.
- 1. A hashed value of image and masked-image combination linked to users for cryptographic image identification.
- 2. A Machine Learning model trained with maskless and masked human image pairs for image recognition purposes
- 3. First an online facial image database is used,
- 4. Several participants are also recruited.



A Masked-image recognition system

- 1. Blockchain model for cost-effective image hash storage and retrieval
- 2. Machine learning and AI recommender system for image recognition
- 3. User Interface Frontend forms and Backend prototype test cases, design, and testing





Smart Sensor for EEG Acquisition and Epileptic Seizure Detection and prediction

With

Prof Justin Dauwels, Nanyang Technological University, Singapore Prof Alok Mishra, Atilim University, Turkey Prof K Ramesh, Karnataka State Women's University, Vijayapura, India Dr Gaurav Garg, Ulster University, UK Students:

Ms Sylvia Bugeja, Mr Eliazar Elisha Audu, Ms. Noela Galea, Ms. Yezi Ali Kadhim, Mr Sean Bugeja, Mr. James Bonello



Epilepsy

- Epilepsy is a medical condition associated with recurrent seizures, which disrupt normal electrical function of the brain as a result of excessive synchronization (hyper-synchronization) of cortical neural network.
- It has profound effects on the state of consciousness, cognitive function and bodily motor control of the affected persons at the onset of seizures.
- Statistical studies from World Health Organization (WHO) revealed that about 1% (or 50 Million) people worldwide are currently affected by epilepsy.





Epileptic Seizure

- In about 70% of epileptic patients, the prognosis for managing seizures is good.
- However, about 25% of the people living with epilepsy are resistant to anti-epilepsy drug treatment or cannot be cured by surgery.
- About 30% of the epileptic patients remain untreated.
- People living with epilepsy faces risk of sudden death, high possibility of emergency room cases, social discrimination and stigmatization, and socio-economic problems.







Epileptic Seizure



Epileptic Seizure Detection and prediction

- Good news!! Interdisciplinary research offers clinical management of epilepsy through the application of signal processing and machine learning techniques.
- Epileptic seizure detection aims to develop systems which monitor patient EEG, learn to classify whether it is seizure or non-seizure EEG and act upon such a decision
- Epileptic seizure prediction aims to develop systems which monitor patient EEG, learn to predict whether the present signal is indicating the provability of occurrence of a seizure in a given time.





Epileptic Seizure Prediction

- Use signal processing and machine learning techniques
 - Extract features -> Create feature space -> Train -> Learn
- Patient-specific vs. Patient-non-specific systems
 - Patient-non-specific systems do not perform well across a large patient population, therefore not practical





Epileptic Seizure Prediction



Epileptic Seizure Prediction



Extracting seizure and non-seizure sets

- Each hour is first segmented into 2-second epochs of EEG
- Extracting non-seizure sets is very simple
 - Divide non-seizure hour into smaller sets of some decide equivalent length (eg. 10 minutes)
 - 2. Store selected sets in patient folder



Extracting seizure and non-seizure sets

- Extracting seizure sets is less trivial
 - May have multiple seizures recorded in one hour of EEG
- Get seizure hour
 - For every seizure in the hour
 - 1. Extract set of appropriate length (eg. 10 minutes) such that no other seizure EEG is contained within the set
 - 2. Store details of acquired seizure set in reference table
 - 3. Store seizure set in patient folder
- Acquires smaller sets that still contain the necessary seizure EEG





Extracting seizure and non-seizure sets



Extracting features

- The power is calculated from the frequency domain by squaring the amplitude
- For every EEG epoch
 - Convert from the time domain to the frequency domain
 - Uses Fourier Transform
 - Take the frequencies over the range {1, 2, 3, ..., 10}
 - Calculate the total power over these frequencies
 - Take the average
- Why the lower frequencies only?
 - Seizures generally act on the lower frequencies of EEG
- So the feature will be the average power over 1-10 Hz

-Università



Medical images: characteristics

• The prominent waves present in EEG recordings



Extracting features

Multi-channel EEG in time domain



Extracting features

• Discrete Wavelet Transform



UNIVERSITY OF MALTA L-Università ta' Malta



Creating the feature space

- Created by concatenating the patient folder sets
- Vary in size across different patients
 - Depending on number of channels and specified set length
- Represents temporal evolution of average power
 - 3-component feature vector containing 3 consecutive epochs
 - The "patient-specific" feature of the training set
- EEG-data
 - Remove artefact channels
 - Transform into temporal representation
 - Add to feature space


Creating the feature space

- Supervised learning approach
- Created by using seizure start and end details from the reference table
- Feature vectors that contain seizure epochs = 1
- Feature vectors that contain \underline{no} seizure epochs = -1





Creating the feature space



Training

• For every patient we now have a feature space and its label vector

- SVM used for training on these structures
- Applies 3-fold cross validation





Software and data used

- Method applied on CHB-MIT EEG datasets (as in Shoeb et. al.)
 - Pediatric EEG
- EEGLab is a toolbox plugin for MATLAB
- EEGLab used for visualisation, manipulation and processing of EEG data
- MATLAB used for building input structures and SVM training





Results: Evaluation Criteria

- Sensitivity
 - percentage of seizure epochs correctly detected

 $Sensitivity = \frac{TP}{TP + FN} \times 100$





Results: Evaluation Criteria

- Sensitivity
 - percentage of seizure epochs correctly detected
- Latency
 - the delay between the actual start of the seizure (or seizure onset) and the time it took the classier to react
- Specificity
 - number of non-seizure epochs falsely classified as seizure epochs

$$Specificity = \frac{TN}{TN + FP} \times 100$$





Results: Evaluation Criteria

- Sensitivity
 - percentage of seizure epochs correctly detected
- Latency
 - the delay between the actual start of the seizure (or seizure onset) and the time it took the classier to react
- Specificity
 - number of non-seizure epochs falsely classified as seizure epochs
- Results are prioritized by sensitivity followed by a tradeoff between latency and false positive number





Results

	FYP	Shoeb et. al.
Sensitivity	92.39%	96%
Latency	3.72 seconds	4.6 seconds
Selectivity	91.55%	
Total number of hours used	49.48 hours	916 hours





Another method

- The proposed method creates a simple, yet very effective training set acquisition for epileptic seizure detection making the classifier's training phase faster.
- The proposed method was tested using CHB-MIT database, a dataset of 977 hours of EEG data containing 192 seizure instances from 22 pediatric patients collected at the Children's Hospital, Boston.





Results

	10-Minute Subsets		20-M Sub	linute osets	30-Minutes Subset		Shoeb et al. ^[2]
	SVM	ELM	SVM	ELM	SVM	ELM	results
Sensitivity(%)	95.33	99.48	95.42	99.48	97.98	98.99	96%
Specificity(%)	87.11	74.21	89.90	77.16	83.73	81.39	-
Latency(Seconds)	3.18	0.97	2.88	0.97	2.95	1.26	3



UNIVERSITY OF MALTA L-Università ta' Malta



Epileptic Seizure Detection using Classifier stacking

With

Simon Xerri



Reference	Year	Dataset	Signal Processing Method	Classification Method	Results
8	2020	CHB-MIT	DWT	SVM	Sensitivity: 97.2% Specificity: 100% Accuracy: 98.6%
[20]	2010	Bonn	DWT and Line Length Feature Extraction	ANN	Sensitivity: 99.4% Specificity: 100% Accuracy: 99.60%
[22]	2011	Bonn	DWT	K-Means Clustering and MLPNN	Sensitivity: 100% Specificity: 98.04% Accuracy: 99.60%
23	2007	Bonn	DWT	SVM	Accuracy: 99.28%
[24]	2019	Bonn	EMD	SVM K-Nearest Neighbour Decision Tree	Accuracy: 100% Accuracy: 100% Accuracy: 96.67%
25	2016	Bonn	DWT	ANN SVM	Accuracy: 93.9% Accuracy: 99.97%
26	2018	CHB-MIT	DWT	SVM	Accuracy: 95.6%
[27]	2020	Bonn	DWT	ANN	Sensitivity: 96% Specificity: 99% Accuracy: 98.4%
29	2011	Bonn	AE	SVM	Sensitivity: 93.75% Specificity: 82.5% Accuracy: 88.0%
32	2011	Bonn	DWT and AE	PNN	Accuracy: 99% - 100%
35	2012	Bonn	EMD	SVM	Sensitivity: 100% Specificity: 99.38% Accuracy: 99.50%
[36]	2011	CHB-MIT	EMD	RF, K-Nearest Neighbour, Functional Tree, and Bayes Net	Sensitivity: 97.91% Specificity: 99.57% Accuracy: 99.41%
37	2013	Bonn	EMD	SVM	Sensitivity: 98.0% Specificity: 99.04%
38	2019	Bonn	EMD	SVM	Accuracy: 97.30%
39	2018	Bonn	EMD	SVM	Accuracy: 100%
[40]	2020	CHB-MIT	EMD	ANN	Sensitivity: 100% Specificity: 97.92% Accuracy: 99.06%
50	2013	CHB-MIT	DWT	ANN	Accuracy: 90.0%
51	2020	CHB-MIT	Channel Selection	SVM	Sensitivity: 96.87% Specificity: 99.95%





The epileptic seizure detection system architecture



Classifier Stacking: an ensemble learning technique



Results: confusion matrix





UNIVERSITY OF MALTA L-Università ta' Malta



Results: Single classifier

Classifier	Accuracy	Sensitivity	Specificity
SVM	0.8875	0.8430	0.9318
NB	0.6838	0.4258	0.9417
KNN	0.9316	0.8948	0.9684
RF	0.9313	0.9033	0.9597
ELM	0.8883	0.8576	0.9190
MLPNN	0.9473	0.9438	0.9514





alta

Results: Two classifiers stacking

Classifiers	Accuracy	Sensitivity	Specificity
SVM + NB	0.8928	0.8796	0.9063
SVM + KNN	0.9374	0.9289	0.9460
SVM + RF	0.9350	0.9253	0.9446
SVM + ELM	0.8965	0.9028	0.8903
SVM + MLPNN	0.9460	0.9423	0.9494
NB + KNN	0.9371	0.9225	0.9516
NB + RF	0.9347	0.9203	0.9490
NB + ELM	0.8901	0.8546	0.9257
NB + MLPNN	0.9392	0.9433	0.9353
KNN + RF	0.9462	0.9364	0.9561
KNN + ELM	0.9309	0.8960	0.9658
KNN + MLPNN	0.9530	0.9520	0.9541
RF + ELM	0.9230	0.9030	0.9568
RF + MLPNN	0.9529	0.9414	0.9646
ELM + MLPNN	0.9474	0.9314	0.9634





versita

ta

Results: Three classifiers stacking

Classifiers	Accuracy	Sensitivity	Specificity
SVM + NB + KNN	0.9386	0.9271	0.9501
SVM + NB + RF	0.9322	0.9216	0.9434
SVM + NB + ELM	0.8943	0.9008	0.8877
SVM + NB + MLPNN	0.9500	0.9500	0.9504
SVM + KNN + RF	0.9477	0.9404	0.9550
SVM + KNN + ELM	0.9297	0.9126	0.9469
SVM + KNN + MLPNN	0.9552	0.9509	0.9596
SVM + RF + ELM	0.9294	0.9127	0.9461
SVM + RF + MLPNN	0.9558	0.9499	0.9617
SVM + ELM + MLPNN	0.9478	0.9459	0.9505
NB + KNN + RF	0.9485	0.9387	0.9584
NB + KNN + ELM	0.9312	0.8986	0.9638
NB + KNN + MLPNN	0.9556	0.9488	0.9624
NB + RF + ELM	0.9309	0.9044	0.9576
NB + RF + MLPNN	0.9546	0.9478	0.9616
NB + ELM + MLPNN	0.9477	0.9413	0.9542
KNN + RF + ELM	0.9357	0.9212	0.9501
KNN + RF + MLPNN	0.9559	0.9512	0.9607
KNN + ELM + MLPNN	0.9468	0.9468	0.9472
RF + ELM + MLPNN	0.9473	0.9394	0.9551

Results: Four classifiers stacking

Classifiers	Accuracy	Sensitivity	Specificity
SVM + NB + KNN + RF	0.9493	0.9431	0.9555
SVM + NB + KNN + ELM	0.9319	0.9132	0.9506
SVM + NB + KNN + MLPNN	0.9538	0.9540	0.9535
SVM + KNN + RF + ELM	0.9361	0.9153	0.9570
SVM + KNN + RF + MLPNN	0.9521	0.9376	0.9669
SVM + KNN + ELM + MLPNN	0.9370	0.9461	0.9283
SVM + RF + ELM + MLPNN	0.9466	0.9390	0.9546
NB + KNN + RF + ELM	0.9372	0.9200	0.9546
NB + KNN + RF + MLPNN	0.9553	0.9453	0.9654
NB + RF + ELM + MLPNN	0.9448	0.9447	0.9454
KNN + RF + ELM + MLPNN	0.9487	0.9286	0.9686





ta

Results: Five classifiers stacking

Classifiers	Accuracy	Sensitivity	Specificity
SVM + NB + KNN + RF + ELM	0.9373	0.9112	0.9638
SVM + NB + KNN + RF + MLPNN	0.9600	0.9535	0.9665
SVM + NB + KNN + MLPNN + ELM	0.9469	0.9463	0.9480
SVM + NB + RF + MLPNN + ELM	0.9450	0.9341	0.9558
NB + KNN + RF + ELM + MLPNN	0.9488	0.9284	0.9694



UNIVERSITY OF MALTA L-Università ta' Malta



lta

Results: Six classifiers stacking				
Classifiers	Accuracy	Sensitivity	Specificity	
SVM + NB + KNN + RF + ELM + MLPNN	0.9500	0.9368	0.9637	





Results: Summary: classifiers stacking

Classifiers	Accuracy	Sensitivity	Specificity
MLPNN	0.9473	0.9438	0.9514
KNN + MLPNN	0.9530	0.9520	0.9637
KNN + RF + MLPNN	0.9559	0.9512	0.9607
NB + KNN + RF + MLPNN	0.9553	0.9453	0.9654
SVM + NB + KNN + RF + MLPNN	0.9600	0.9535	0.9665
SVM + NB + KNN + RF + ELM + MLPNN	0.9500	0.9368	0.9637



UNIVERSITY OF MALTA L-Università ta' Malta



Epileptic Seizure localization

- If we can detect or predict seizure onsets by using the less number of channels (ideally only one)
- It will help us in making the seizure detection and prediction energy efficient.





With

Dr. Gaurav Garg

Prof. Girijesh Prasad

Dr. Damien Coyle

Aim:

To provide novel methodologies to identify the biomarkers for neurodegenerative diseases using the functional neuroimaging data.

ta' Malta

Neurodegenerative disease

Neurodegeneration is a term for the progressive loss of structure or function of neurons, including death of neurons.

Examples of neurodegenerative diseases include

- Amyotrophic lateral sclerosis (ALS),
- Parkinson's disease and
- Alzheimer's disease.



Alzheimer's Disease

- The damage initially appears to take place in the hippocampus, the part of the brain essential in forming memories.
- As more neurons die, additional parts of the brain are affected.
- By the final stage of Alzheimer's, damage is widespread, and brain tissue gets shrunk significantly.





Alzheimer's Disease



Source: <u>NIH</u>



UNIVERSITY OF MALTA L-Università ta' Malta





Functional Neuroimaging

Subcortical microelectrode implant and high-field fMRI.

- LFP: Local field potential
- BOLD: Blood oxygen level dependent
- CBV: Cerebral blood volume
- CBF: Cerebral blood flow



Alzheimer's Disease



Illustration of fMRI of human brain for an auditory paradigm (Garg et. al., 2011)





Alzheimer's Disease



Most significant regions in AD as suggested in state-of-the-art literature.



UNIVERSITY OF MALTA L-Università ta' Malta



Alzheimer's Disease

YH-EH freq 0.01Hz to 0.167Hz



YH-AD freq 0.01Hz to 0.167Hz



EH-AD freq 0.01Hz to 0.167Hz



Significant differences based on frequency of metabolism based analysis of resting state fMRI data. These regional differences may be considered as biomarkers for the underlying pathology. Here, blue represent decrease while yellow represent increase in frequency.

UNIVERSITY OF GALTA al., 2015

L-Università ta' Malt

Dealing with the noise in the data

Existing limitations

• Environmental noises near cistern regions or air cavities are widely left untreated, and distortion of the recorded functional data poses diminishing effects on the analysis.

Proposed Solution

- Classify the data in terms of several Gaussian processes.
- Replace the values of the data from the mean of the identified Gaussian process. This step removes the distortion caused due to medium dissimilarity and fading of the signals near the boundary regions Università

Alzheimer's Disease



Brain fMRI slice with AWGN: for untreated data (a) Reference template

(b) Results; and similarly for data treated with GMM:

(c) Reference template and

(d) Results for SNR of 16, 8, 4, 2, 1 and 0 dBs respectively

from left to right.

-Università



Estimation of stimulus-driven activation

Existing limitations

• State-of-the-art time domain analysis is based on a model definition where the accuracy of the results depends on the accuracy of the defined model. While, the data driven methods mostly based on independent component analysis (ICA) may give accurate results, but these methods often needs human intervention to decide about the appropriate component to select or discard.

Proposed Solution

• A hybrid method which includes the benefits of model driven and data-driven analysis has been developed. This method is helpful in automating the estimation process in the time domain analysis.





Estimation of stimulus-driven activation



Estimating pace of the metabolic activity

Existing limitations

• Existing resting state analysis methods take account of the total power in the given time-series. Though, this strategy is useful to identify the strength of the functional activity, but lack the ability to distinguish the relative pace such as slowing or increasing the metabolism of the brain, wherein, the pace of the neuronal activity is an important factor in neurodegenerative diseases.

Proposed Solution

• A method, namely regional optimal frequency analysis (ROFA) has been developed to identify the frequency supported by the majority of the voxels at a particular region of the brain.




Estimating pace of the metabolic activity

Regional optimal frequency analysis (ROFA)



Name of the Brain Regions* defined in Automatic

Anatomic Labelling (AAL) Template

Early effects of degeneration on neuronal entities

Existing limitations

• Initial stages of disease don't show any structural change in the brain, but functional changes are evident. These functional changes include task reallocation and overloading of the remaining neurons as a compensatory effect. However, none of the literature has tried to identify these functional changes as metabolically active neuronal assemblies and degenerated neuronal assemblies as recorded at the fMRI voxels.

Proposed Solution

• Cluster movement analysis (CMA) method has been proposed where metabolically active and inactive voxels has been identified based on the BOLD fMRI responses derived from each of the boxes time-series.

Early effects of degeneration on neuronal entities

Cluster Movement Analysis (CMA)





Early effects of degeneration on neuronal entities

Cluster Movement Analysis (CMA)

Superior frontal gyrus, dorsolateral 2 Middle frontal gyrus Count of BOLD intensity change Mean Intensity value change Inferior frontal gyrus, opercular part 3 4 Inferior frontal gyrus, triangular part 40 5 Rolandic operculum 6 Supplementary motor area Superior frontal gyrus, medial 35 8 Cuneus 9 Lingual gyrus 10 Superior occipital gyrus Middle occipital gyrus (MOG) 11 30 Inferior occipital gyrus (IOG) 12 13 Fusiform gyrus (FFG) 14 Superior parietal gyrus 25 15 Inferior parietal 16 Supramarginal gyrus 17 Angular gyrus 20Precuneus 18 19 Paracentral lobule Superior temporal gyrus 20 21 Middle temporal gyrus 15 22 Inferior temporal gyrus .01 .05 .09 .113 17 23. Superior frontal gyrus, orbital part (SFG_OP) 24. Middle frontal gyrus, orbital part Number of the Brain regions from AAL template 25 Inferior frontal gyrus, orbital part Superior frontal gyrus, medial orbital (SFG_MO) 26 27 Gyrus rectus 28 Insula (INS) •YH -----AD 29 Anterior cingulate and paracingulate gyri Median cingulate and paracingulate gyri 30 31 Posterior cingulate gyrus (PCG) 32 Parahippocampal gyrus (PHG) 33 Temporal pole: superior temporal gyrus 34 Temporal pole: middle temporal gyrus 35 Olfactory cortex 36 Hippocampus (HIP) 37 Amygdala (AMYG) 38 Caudate nucleus 39 Lenticular nucleus, putamen 40 Lenticular nucleus, pallidum (LN P) 41 Thalamus 42 Precentral gyrus 43 Calcarine fissure and surrounding cortex Postcentral gyrus 44 45 Heschl gyrus (HESG) Cerebellum UNIVERSITY OF MALTA Vermis-cerebelli L-Università ta' Malta

Part of Cerebrum

Name of the Brain Regions* defined in Automatic

Anatomic Labelling (AAL) Template

Regions Name

No.

Biomarkers for Alzheimer's disease using functional MRI data

Novelty

- A method to remove noises from the cistern regions of the brain
- A method to accurately identify the stimulus dependent activation of the brain in a fully automated system.
- Identifying the pace of the BOLD metabolic activity in a frequency domain framework for resting state data.
- Identifying the early functional changes in the brain along with metabolically active and degenerated neuronal assemblies.



versita

Analysing Reddit Data for the Prediction, and Detection of Depression

With Mr Michael Philip Orenda



Analysing Reddit Data for the Prediction, and Detection of Depression

- Psychological disorders are a growing concern in the world around us. Depression and anxiety are two examples of such disorders which may affect anyone.
- In 2017, more than 970.8 million people around the world suffered from a type of psychological disorder, and almost another 337 million people had a psychological disorder starting to develop.
- With such high numbers, detection of such disorders at an early stage is of utmost importance.





Analysing Reddit Data for the Prediction, and Detection of Depression

- This research analyses different computer models which can be used to predict the depressive disorder in humans.
- Various classifiers are explored and compared with each other to see the pros and cons of each of the classifiers.
- The final results can then be used to determine which is the best classifier or model to apply in the specific scenario based on the requirements of the user.
- This knowledge can then be utilised by a medical professionals, mainly psychologists, as a tool to help them determine the psychological well-being of their patients when it comes to depression.





Analysing Reddit Data for the Prediction, and Detection of Depression

- This research analyses different computer models which can be used to predict the depressive disorder in humans.
- Various classifiers are explored and compared with each other to see the pros and cons of each of the classifiers.
- The final results can then be used to determine which is the best classifier or model to apply in the specific scenario based on the requirements of the user.
- This knowledge can then be utilised by a medical professionals, mainly psychologists, as a tool to help them determine the psychological well-being of their patients when it comes to depression.





Dataset

- 1. Made up of reddit posts and reddit comments. IT consists of two types of users, depressed users and control users.
- For a user to be classified as depressed, s/he had to explicitly state in their posts or comments that they have been diagnosed with depression.
- In total, it contains 214 depressed users who had a total of 90,222 submissions, and 1,493 control users who had a total of 986,360 submissions.





Dataset

Table 2: Table showing statistics of the dataset [33].							
	Train		Test				
	Depressed	Control	Depressed	Control			
Number of users	135	752	79	741			
Number of submissions	49,557	481,837	40,665	504,523			
Average number of submissions per user	367.1	640.7	514.7	680.9			
Average number of days from first submission to last submission	586.43	640.7	514.7	702.5			
Average number of words per submission	27.4	21.8	27.6	23.7			







Cross validation



Cross validation



Evaluation

•
$$Precision = \frac{TP}{TP+FP}$$

• Accuracy =
$$\frac{TP+TN}{TP+TN+FP+FN}$$

•
$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

• Sensitivity =
$$Recall = \frac{TP}{TP+FN}$$

• Specificity =
$$\frac{TN}{TN+FP}$$





Results

Table 7: Evaluation Metrics for the Implemented Models

Model	Precision	Accuracy	Sensitivity	Specificity	F1-Score
Logistic	0.64	0.90	0.46	0.96	0.54
Regression					
Random Forest	0.29	0.76	0.64	0.78	0.40
SVM	0.58	0.90	0.63	0.94	0.60
CNN	0.43	0.86	0.43	0.92	0.43





Emotional Testing on Facebook's User Experience

With

Mr Roberto Stefano Mangion



Emotional Testing on Facebook's User Experience: Aims

- Acquire insight on how user-friendly the product ** is for new users when compared to frequent users.
- Discern which parts of the brain had the largest ** difference between groups and motives behind an individual's emotional state.



- Università ta' Mai



Emotional Testing on Facebook's User Experience: Aims

- Understanding how user's emotions fluctuate when undertaking certain tasks on a software product.
- Exploring the difference in usability aspect of
 Facebook, from new to frequent Facebook users.





Emotional Testing on Facebook's User Experience: Design

SEGMENTATION OF EACH TASK



Emotional Testing on Facebook's User Experience: Design EEG DATA PROCESSING DESIGN



Emotional Testing on Facebook's User Experience: Design

INDIVIDUAL CONDUCTING TASK ON FACEBOOK



Emotional Testing on Facebook's User Experience: Design

• SYSTEM DESIGN FOR 9 NEW AND 9 FREQIENT USERS



Emotional Testing on Facebook's User Experience: Results





UNIVERSITY OF MALTA L-Università ta' Malta



Emotional Testing on Facebook's User Experience: Results





UNIVERSITY OF MALTA L-Università ta' Malta



Emotional Testing on Facebook's User Experience: Conclusion

- Statistical difference between new and frequent Facebook users' alpha and beta frequencies.
- The temporal region, which consists of the limbic system, which is responsible for an individual's mental state of mind, had the most notable difference in beta frequency.





Emotional Testing on Facebook's User Experience: More details



Received February 15, 2020, accepted March 2, 2020, date of publication March 17, 2020, date of current version April 6, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2981418

Emotional Testing on Facebook's User Experience

ROBERTO STEFANO MANGION¹, LALIT GARG¹⁰, (Member, IEEE), GAURAV GARG², (Member, IEEE), AND OWEN FALZON³

¹Department of Computer Information Systems, University of Malta, 2080 Msida, Malta ²ABV-Indian Institute of Information Technology and Management, Gwalior 474015, India ³Centre for Biomedical Cybernetics, University of Malta, 2080 Msida, Malta

Corresponding author: Lalit Garg (lalit.garg@online.liverpool.ac.uk)

ABSTRACT This study aims at understanding how a user's emotions fluctuate when undertaking certain tasks on a social media platform such as Facebook or other software products which may have emotional effects on its user. Specifically, we explored the difference in the usability aspect of Facebook concerning

- Garg L, Dauwels J, Earnest A, Pang L (2013). Tensor based methods for handling missing data in quality-of-life questionnaires. IEEE Journal of Biomedical and Health Informatics. In press. doi: 10.1109/JBHI.2013.2288803. Impact Factor: 1.978, Journal H-Index(SJR:1996-2012): 55, SJR(2012):0.877.
- Garg L, McClean SI, Meenan BJ, Barton M, Fullerton K (2012). Intelligent patient management and resource planning for complex, heterogeneous, and stochastic healthcare systems. In press. IEEE Transactions on Systems, Man, and Cybernetics--Part A: Systems and Humans 42(6): 1332 1345. Impact Factor: 2.093, Journal H-Index(SJR:1996-2011): 54, SJR(2011):0.085, Scopus SNIP(2011):3.594, Scopus citations: 2, ISI Web of Knowledge Citations: 1, scHolar index: 2.
- Garg L, McClean SI, Meenan BJ, Barton M, Fullerton K (2011). An Extended Mixture Distribution Survival Tree for Patient Pathway Prognostication. Communications in Statistics: Theory and Methodology. 42(16):2912-2934. Impact Factor: 0.351, Journal H-Index(SJR:1996-2011): 24, SJR(2011):0.035, Scopus SNIP(2011):0.550.
- Barton M, McClean SI, Gillespie J, Garg L, Wilson D, Fullerton K (2012). Is it beneficial to increase the provision of thrombolysis?- A discrete-event simulation model, QJM: An International Journal of Medicine. 105(7), (art. no. hcs036):665-673. Impact Factor: 2.146, Journal H-Index(SJR:1993-2011): 71, SJR(2011):0.18, Scopus SNIP(2011):0.897, Scopus citations: 3, ISI Web of Knowledge Citations: 1, scHolar index: 2.
- Garg L, McClean SI, Meenan BJ, Millard PH (2011). Phase-type survival trees and mixed distribution survival trees for clustering patients' hospital length of stay. INFORMATICA. 22(1): 57-72. Impact Factor: 1.786, Journal H-Index(SJR:1996-2010): 15, SJR(2011):0.048, Scopus SNIP(2011):1.814, Scopus citations: 8, ISI Web of Knowledge Citations: 7, scHolar index: 10.





- McClean SI, Barton M, Garg L, Fullerton K (2011). A Modeling framework that combine Markov models and discrete event simulation for stroke patient care. ACM Transactions on Modelling and Computer Simulation (TOMACS), 21(4). Impact Factor: 0.684, Journal H-Index(SJR:1991-2009): 26, SJR(2011): 0.043, Scopus SNIP(2011): 2.053, ISI Web of Knowledge Citations: 2, Scopus citations: 4, scHolar citations: 6.
- Jahangirian M, Eldabi T, Garg L, Jun T, Naseer A, Patel B, Stergioulas L, Young TP (2011). A Rapid Review Method For Extremely Large Corpora of Literature: Applications to the domains of Modelling, Simulation, and Management. International Journal of Information Management (IJIM), 31(3):234-243. Impact Factor: 1.554, Journal H-Index(SJR:1986-2010): 32,SJR(2011):0.043, Scopus SNIP(2011):2.517, Scopus citations: 3,ISI Web of Knowledge Citations: 2, scHolar citations: 5.
- Gillespie J, McClean SI Scotney B, Garg L, Barton M, Fullerton K (2011). Costing Hospital Resources for Stroke Patients using Phase-type Models. Health Care Management Science. 14(3):279-291. Impact Factor: 0.394, SJI(2011): 0.063, Scopus SNIP(2011): 1.169, Journal H-Index(SJR:1998-2009): 23, Scopus citations: 1, scHolar citations: 3.
- Papadapoulou A, Tsaklidis G, McClean SI, Garg L (2011). On the moments and the distribution of the cost of a semi Markov model for healthcare systems. Methodology and Computing in Applied Probability.14(3):717-737. Impact Factor: 0.774, Journal H-Index(SJR:2004-2010): 11, SJR(2011):0.040, Scopus SNIP(2011):0.718.
- Garg G, Prasad G, Garg L, Coyle D (2011). Gaussian Mixture Models for Brain Activation Detection from fMRI Data, International Journal of Bioelectromagnetism. 13(4):255-260, scHolar citations: 2.





- Garg L, McClean SI, Meenan BJ, Millard PH (2010). A non-homogeneous discrete time Markov model for admission scheduling and resource planning in a care system. Health Care Management Science. 13(2):155–169. Impact Factor: 0.394, Journal H-Index(SJR:1998-2009): 23, SJI(2011):0.063, Scopus SNIP(2011):1.169, Scopus citations: 8, ISI Web of Knowledge Citations: 5, scHolar citations: 18.
- Garg L., McClean SI, Barton M, Meenan BJ and Fullerton K (2010). The extended mixture distribution survival tree based analysis for clustering and patient pathway prognostication in a stroke care unit. International Journal of Information Sciences and Application. 2(4): 671-675.
- Garg L, McClean SI, Meenan BJ, Millard PH (2009). Non-homogeneous Markov Models for Sequential Pattern Mining of Healthcare Data. IMA journal Management Mathematics. 20(4): 327-344. Journal H-Index(SJR:2001-2009): 11,SJR(2011):0.033, Scopus SNIP(2011):0.671, Scopus citations: 1, ISI Web of Knowledge Citations: 2, scHolar citations: 4.
- Garg L, McClean SI, Barton M (2008). Is Management Science doing Enough to Improve Healthcare?. International Journal of Human and Social Sciences 3(2): 99-103. scHolar citations: 4.
- Garg L, McClean SI, Meenan BJ, Millard PH (2009). A phase type survival tree model for clustering patients according to their hospital length of stay. Proceedings of the XIII International Conference on Applied Stochastic Models and Data Analysis (ASMDA 2009), pp. 497-502, Eds: Leonidas Sakalauskas, Christos Skiadas, Edmundas K. Zavadskas, ISBN: 978-9955-28-463-5, Publisher: Vilnius Gediminas Technical University Press. scHolar citations: 3.





- Barton M, McClean SI, Garg L, Fullerton K (2009). Modelling Stroke Patient Pathways using Survival Analysis and Simulation Modelling. Proceedings of the XIII International Conference on Applied Stochastic Models and Data Analysis (ASMDA 2009), pp. 370-373, Eds: Leonidas Sakalauskas, Christos Skiadas, Edmundas K. Zavadskas, ISBN: 978-9955-28-463-5, Publisher: Vilnius Gediminas Technical University Press. scHolarcitations: 4.
- McClean SI, Garg L, Barton M, Fullerton K, Millard PH (2009). Using Markov Systems to plan Stroke Services. In Intelligent Patient Management (Eds. McClean, S.; Millard, P.; El-Darzi, E.; Nugent, C.D.), Book Series: Studies in Computational Intelligence, Vol. 189. pp. 241-256. Journal H-Index(SJR:1998-2010): 10, SJI(2011):0.029, Scopus SNIP(2011):0.310, Scopus citations: 2, ISI Web of Knowledge Citations: 4, scHolar citations: 3.
- McClean SI, Millard PH, Garg L (2008). Using Markov Models for Decision Support in Management of High Occupancy Hospital Care. In Intelligent Techniques and Tools for Novel System Architectures (P. Chountas, I. Petrounias, J. Kacprzyk Eds), Springer's book series Studies in Computational Intelligence (SCI). Vol. 109/2008. pp. 187–199. Journal H-Index(SJR:1998-2010): 10, SJI(2011):0.029, Scopus SNIP(2011):0.310.
- McClean SI, Garg L, Meenan BJ, Millard PH. (2007). Non-Homogeneous Markov Models for Performance Monitoring in Healthcare. (In C.H. Skiadas, Eds.) Recent Advances In Stochastic Modelling and Data Analysis. World Scientific Publishing, pp. 146-153. scHolar citations: 4.



