Digital Twin and Smart Health: Shaping the future of healthcare

Digital Twin and Smart Health: Shaping the future of healthcare

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

ediTec

L-Università ta' Malta

Global health challenges

Global health challenges

1. Obesity and Chronic diseases
2. Aging
3. Drug resistance, Hospital acquired infections and medical errors
4. Global warming and pollution
5. Health inequality and healthcare finance
6. Infectious and/or zoonotic diseases and viruses
7. Stress and sleep aponia
8. Relationships and social health

MediTec

L-Università ta' Malta

Future healthcare technologies

ediTec

L-Università ta' Malta

Future of healthcare

20/04/2022

Future healthcare technologies

1. AI/ML: predictive/prescriptive analytics and digital twin
2. Sensors: Smart wearables, cyborg, satellites, wireless sensor networks and IoTs
3. 5G/6G: Cloud computing, telemedicine and Mobile health
4. Robotics (computer vision and natural language processing)
5. Gene editing, genomics, epigenomics proteomics and metabolomic

Enabler: Globalization and economic growth

ediTec

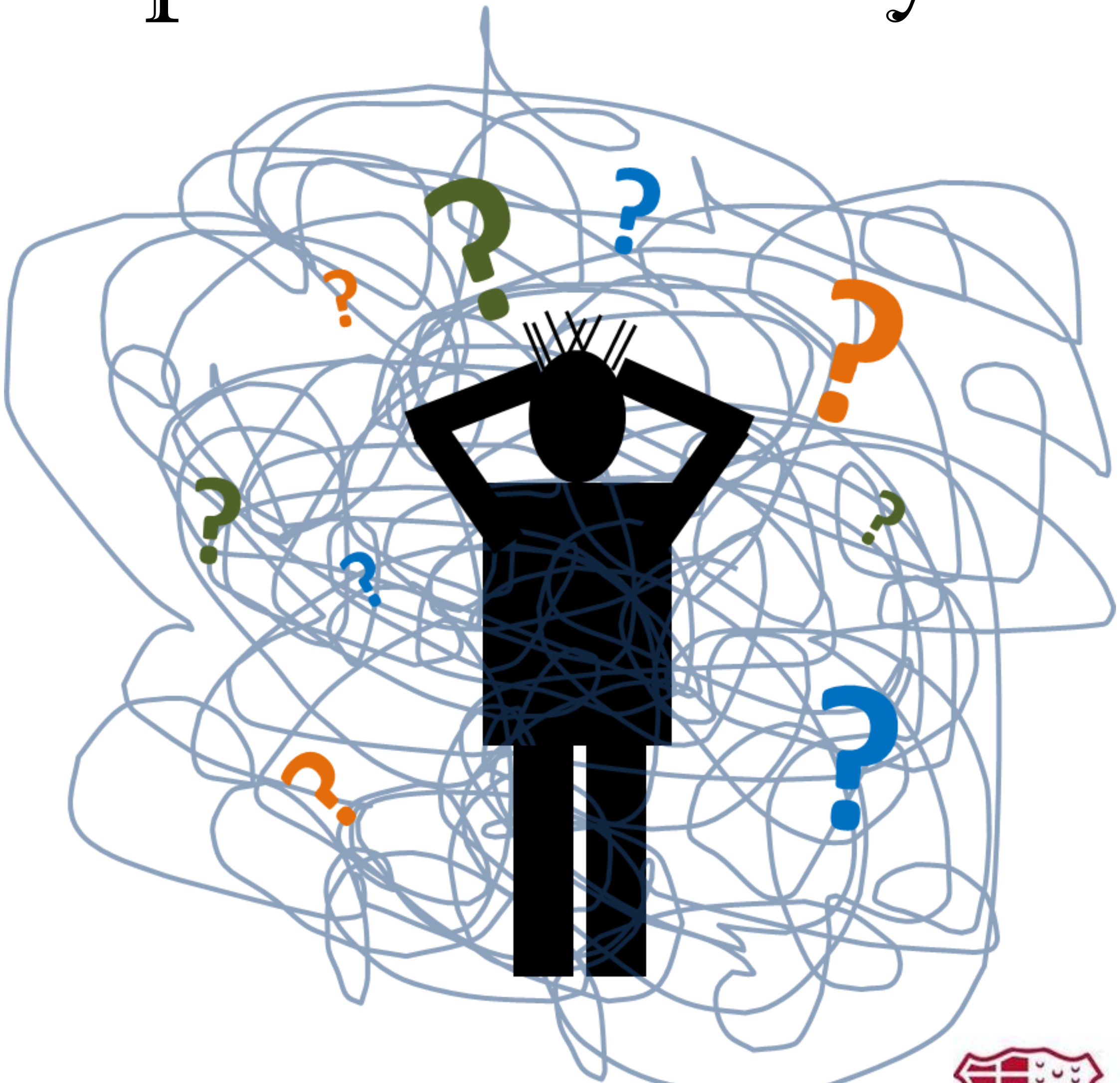
L-Università ta' Malta

Our healthcare system:
A complex system

ediTec

L-Università ta' Malta

Complex Health Systems



20/04/2022

Complex healthcare systems

Holistic approach to healthcare should not just consider medicine but also consider

1. Social aspects
2. Political aspects
3. Cultural aspects
4. Community aspects
5. Communication aspects
6. Transportation aspects

Complex healthcare systems

7. Management aspects
8. Supply chain aspects
9. Administration aspects
10. Education aspects
11. Financial aspects
12. Economical aspects
13. Behavioural/ psychological aspects
14. Geological aspects

ediTec

L-Università ta' Malta

Smart health

Smart Health

Application of artificial intelligence to make the intelligent/right decisions at the right time, right place utilizing the right data. It includes:

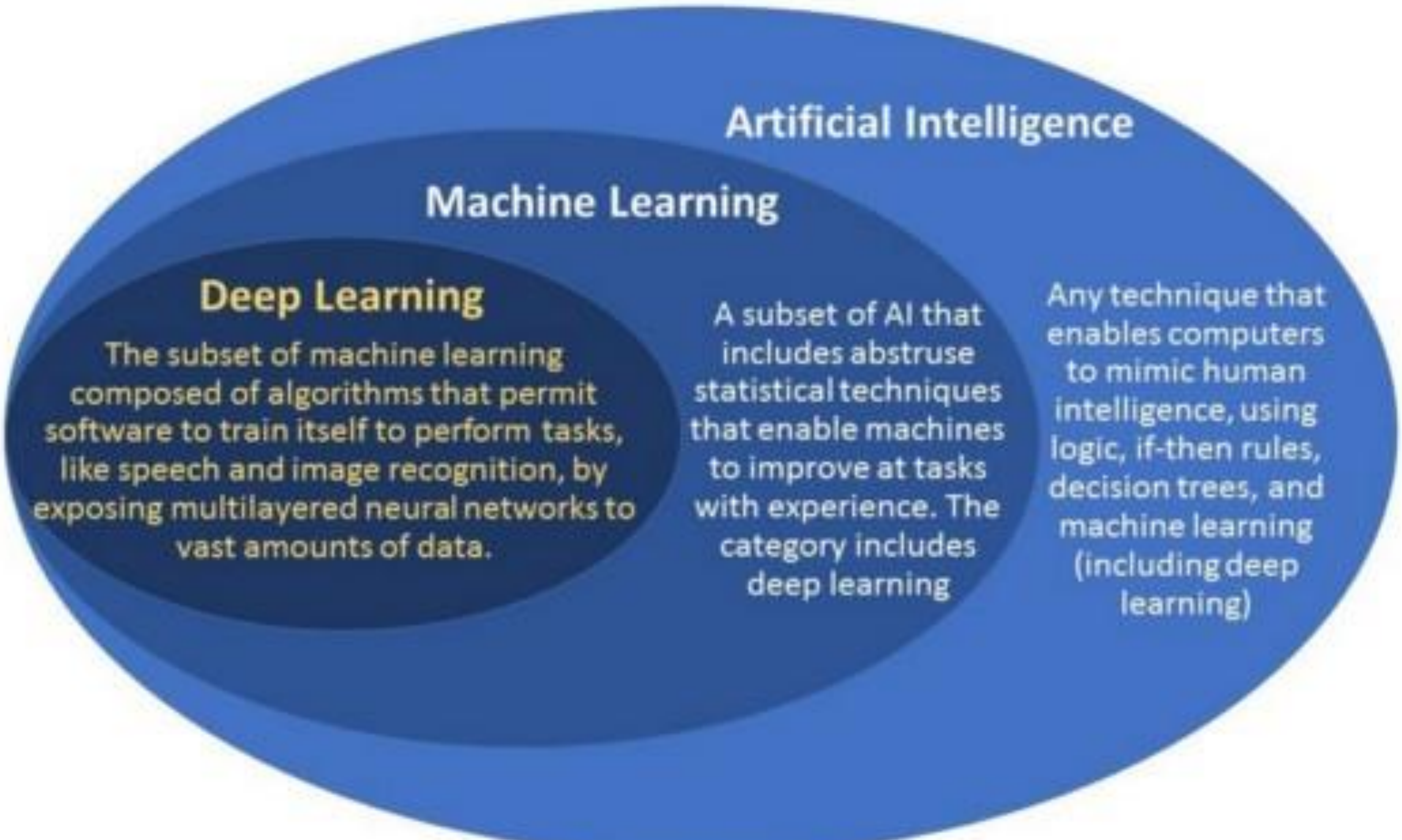
1. Information collection through Sensors and records: Smart wearables, cyborg, satellites, wireless sensor networks and IoTs
2. Information sharing and exchange through 5G/6G: Cloud computing, telemedicine and Mobile health
3. Information processing and decision making: through AI/ML: predictive/prescriptive analytics and digital twin
4. Smart response/action/control through robotics (computer vision and natural language processing)

ediTec

L-Università ta' Malta

Artificial Intelligence

Artificial Intelligence

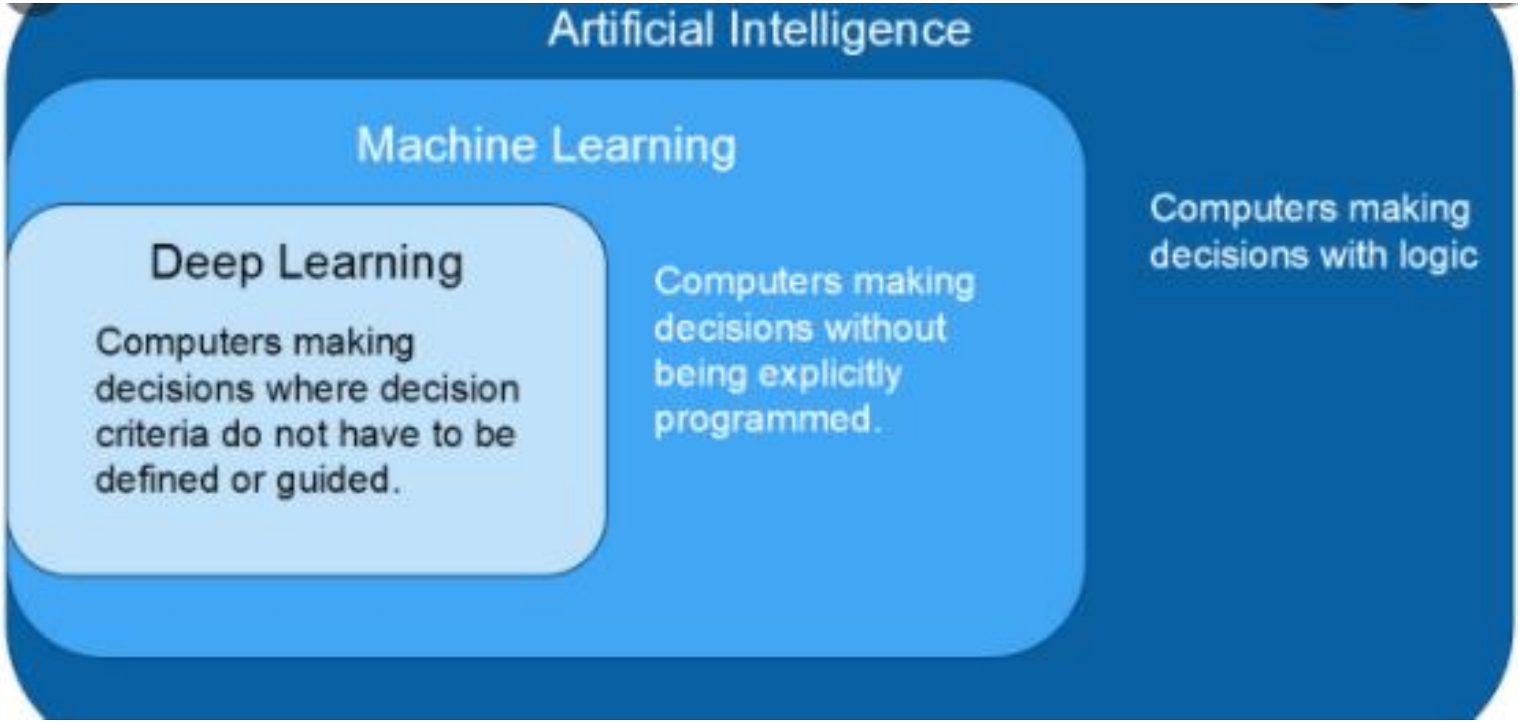


https://medium.com/@harish_6956/what-is-machine-learning-deep-learning-7788604004da

ediTec

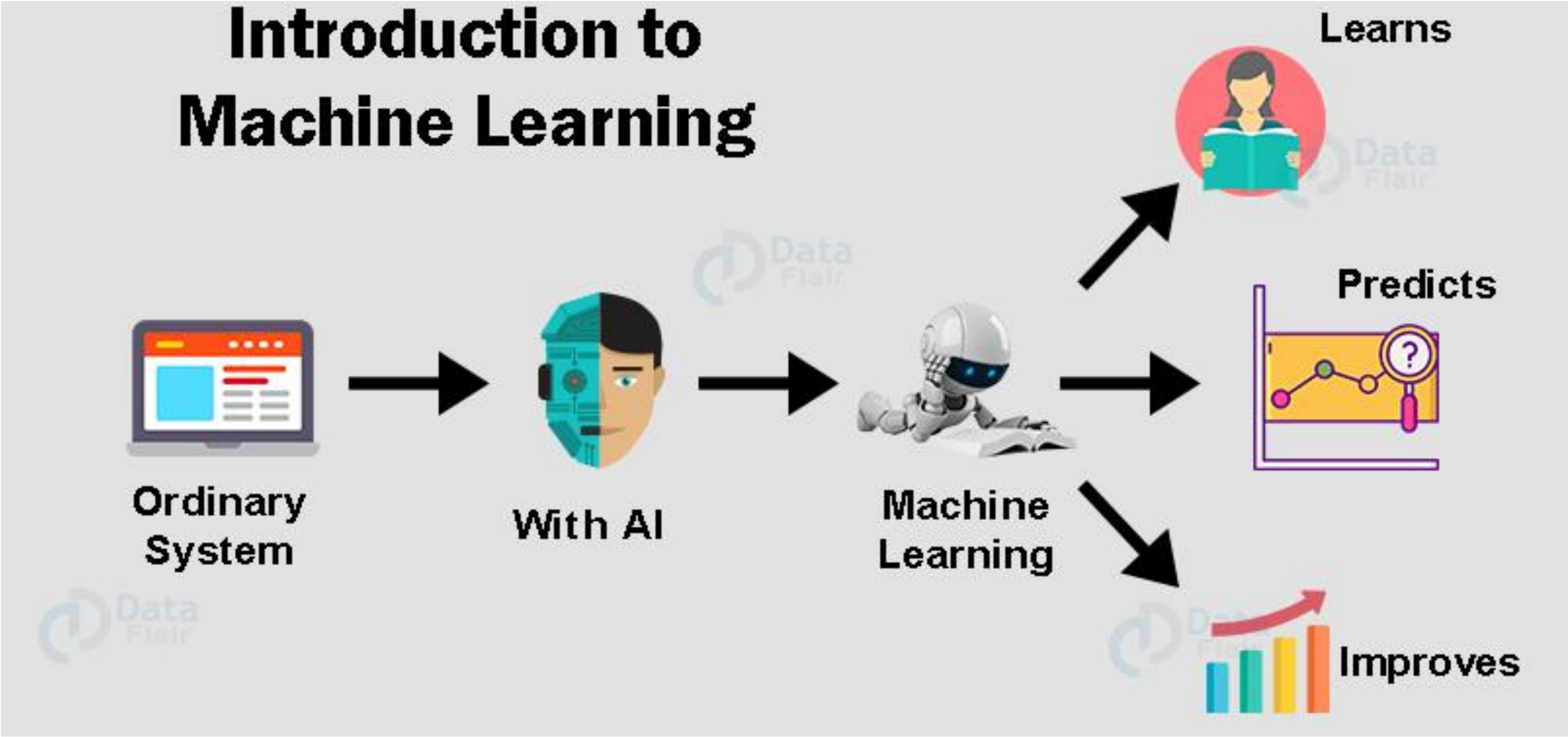
L-Università ta' Malta

Artificial Intelligence



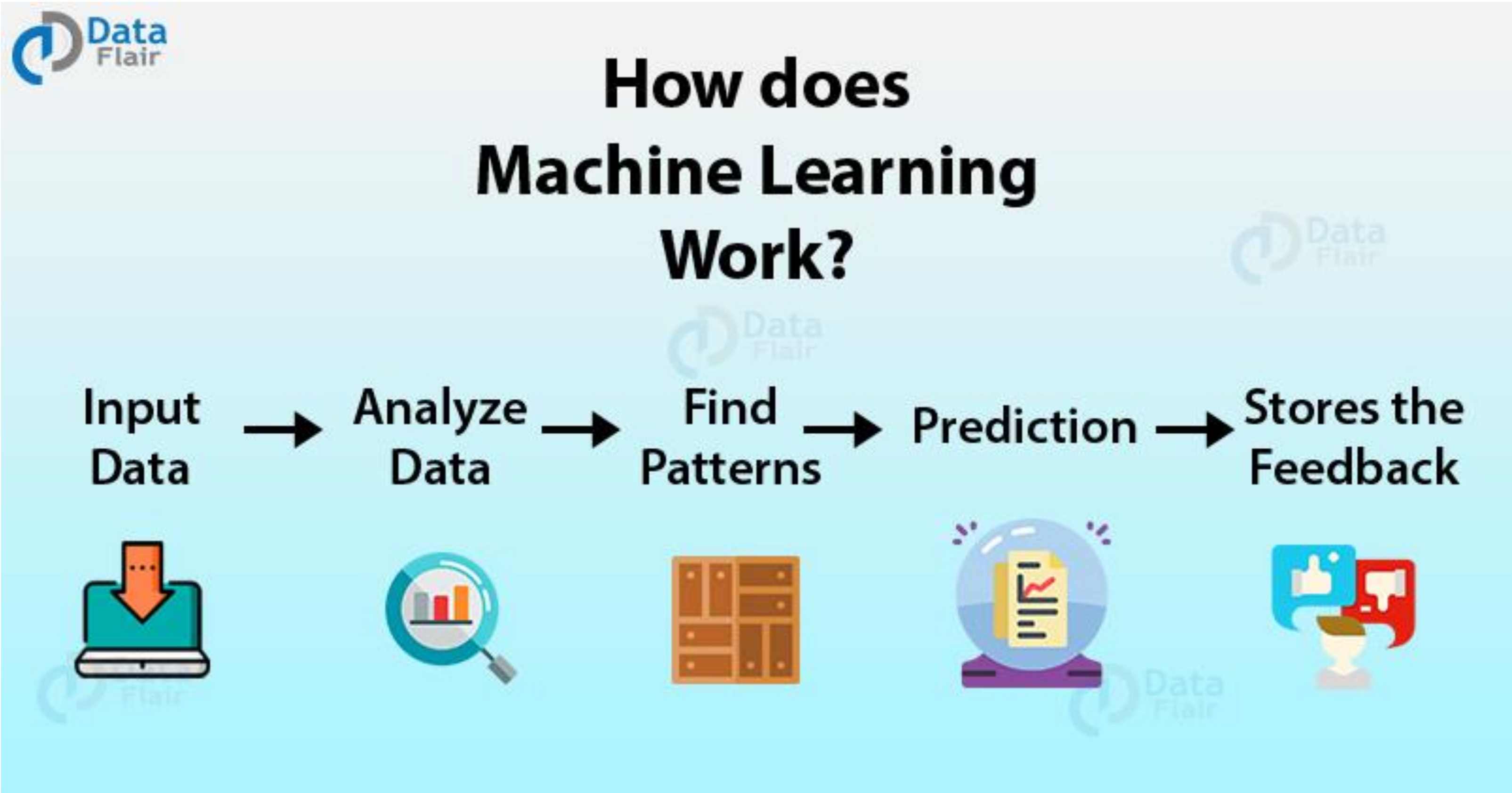
https://medium.com/@harish_6956/what-is-machine-learning-deep-learning-7788604004da

Machine Learning



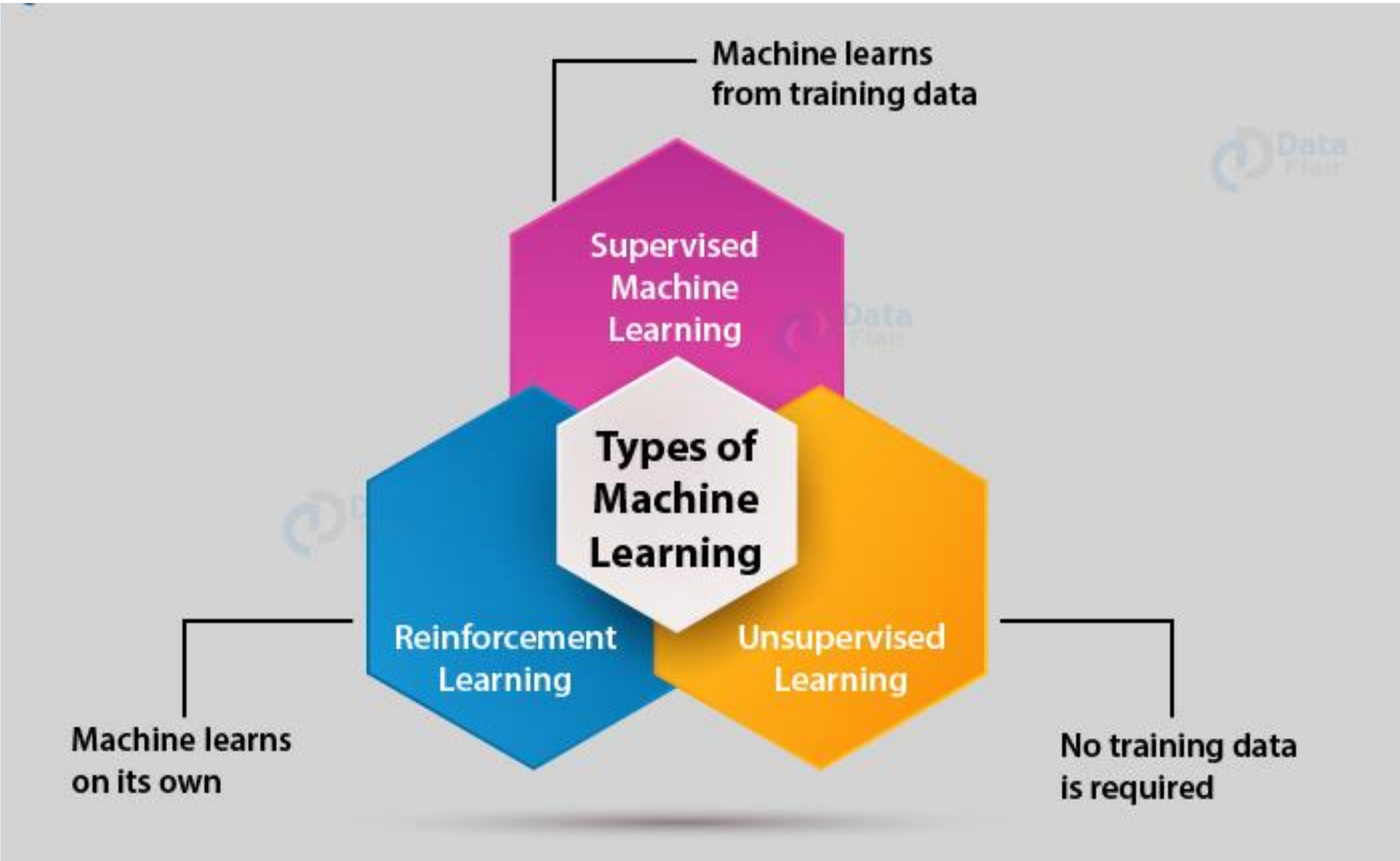
https://data-flair.training/blogs/machine-learning-tutorial/

Machine Learning

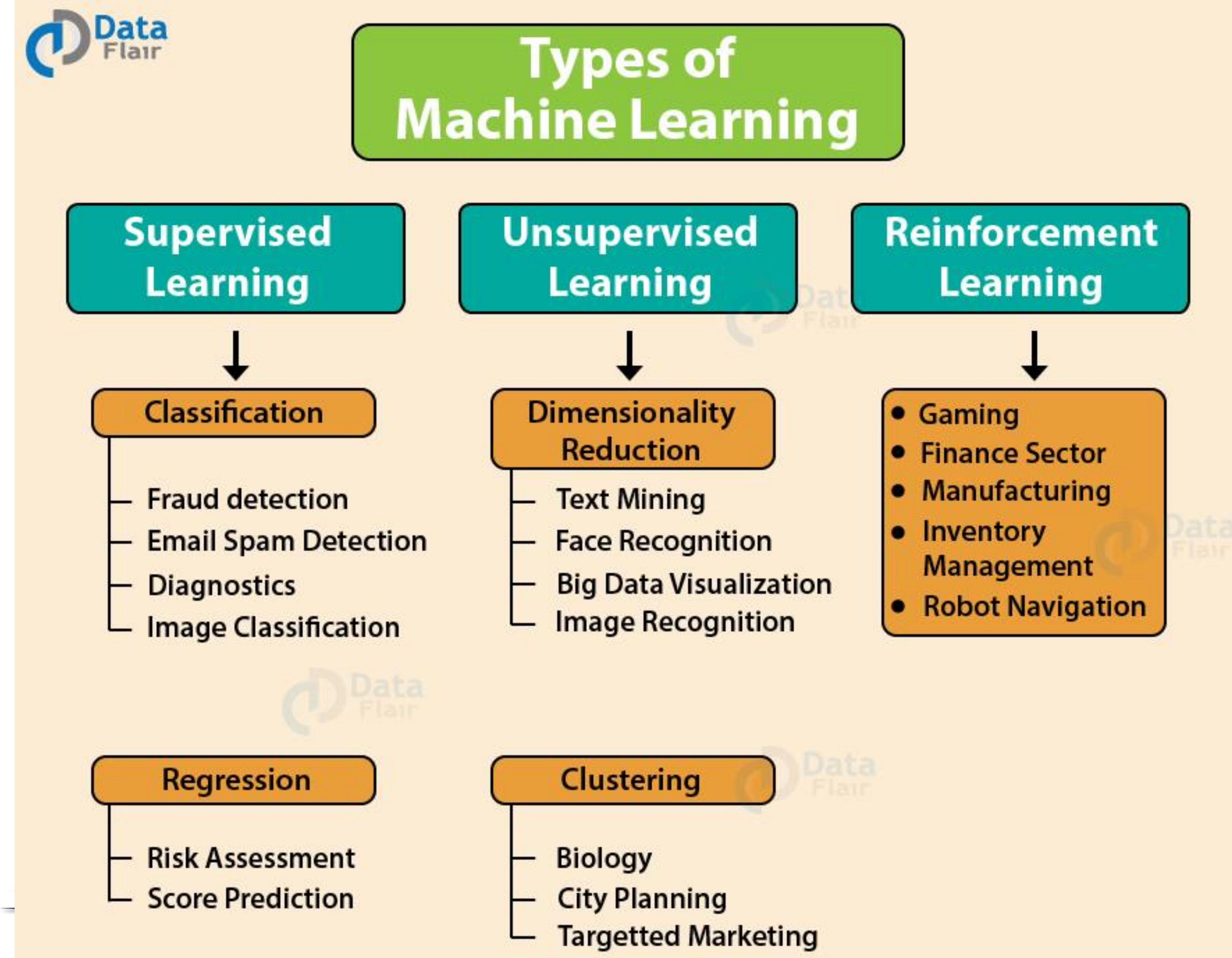


https://data-flair.training/blogs/machine-learning-tutorial/

Machine Learning



https://data-flair.training/blogs/machine-learning-tutorial/


Data
Flair
Types of
Machine Learning
Supervised
Learning
Unsupervised
Learning
Reinforcement
Learning
Classification
Fraud detection
Email Spam Detection
Diagnostics
Image Classification
Regression
Risk Assessment
Score Prediction
Dimensionality
Reduction
Text Mining
Face Recognition
Big Data Visualization
Image Recognition
Clustering
Biology
City Planning
Targetted Marketing
• Gaming
• Finance Sector
• Manufacturing
• Inventory
Management
• Robot Navigation

à

Predictive and prescriptive analysis in healthcare

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

Predictive analysis in healthcare

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

Predictive analysis in healthcare

Predictive analysis in healthcare

Digital twin in healthcare

Digital twin in healthcare

Digital twin in healthcare

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

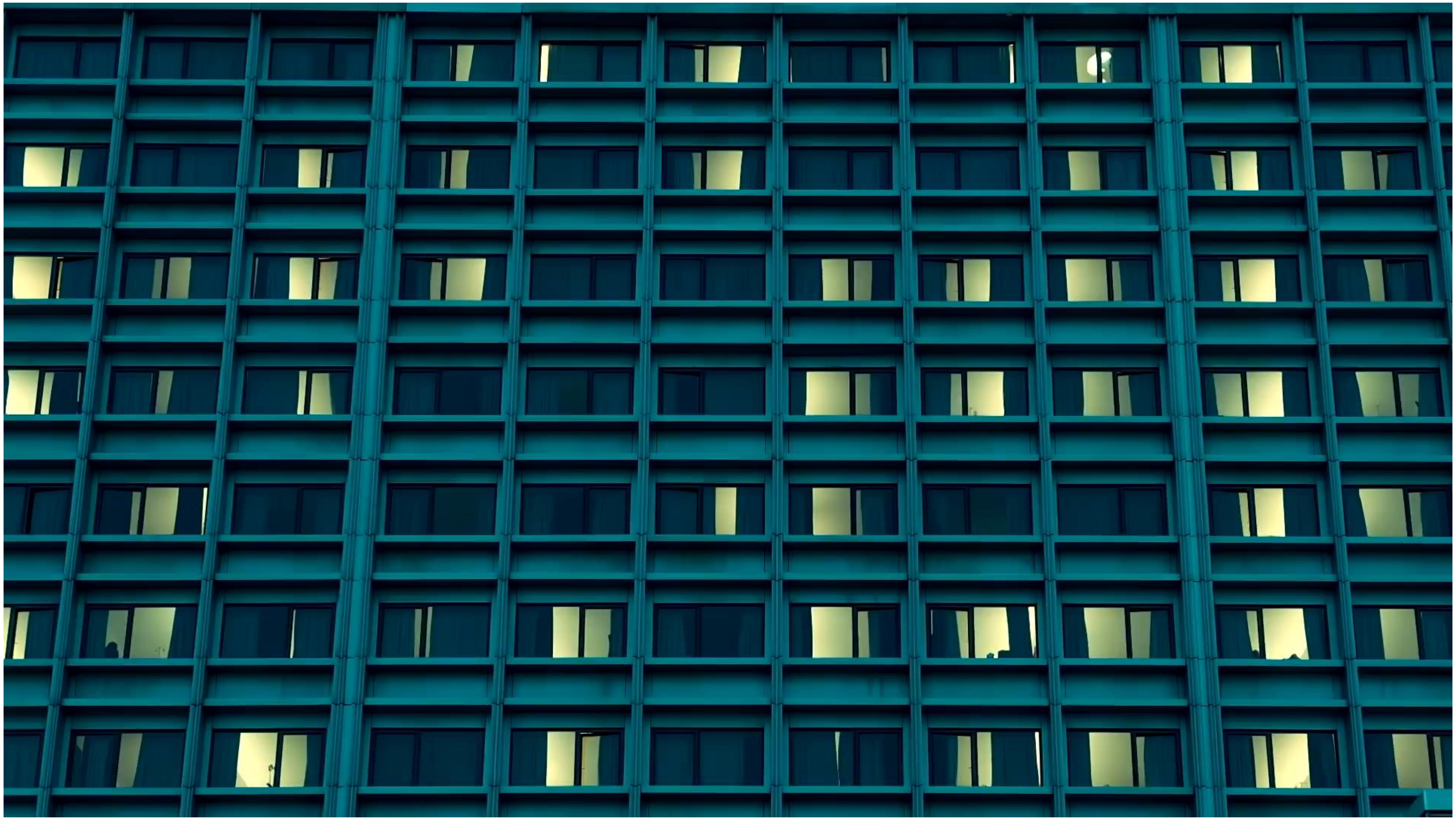
1. Your general information (hight, weight, shape, BMI etc.)
2. Your genomic profile and phenotype
3. Your medical history
4. Your family history (health related)
5. Your diet, activity, travel history
6. Your predictive/prescriptive analysis

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

ediTec

L-Università ta' Malta

Digital twin in healthcare



Digital twin in healthcare

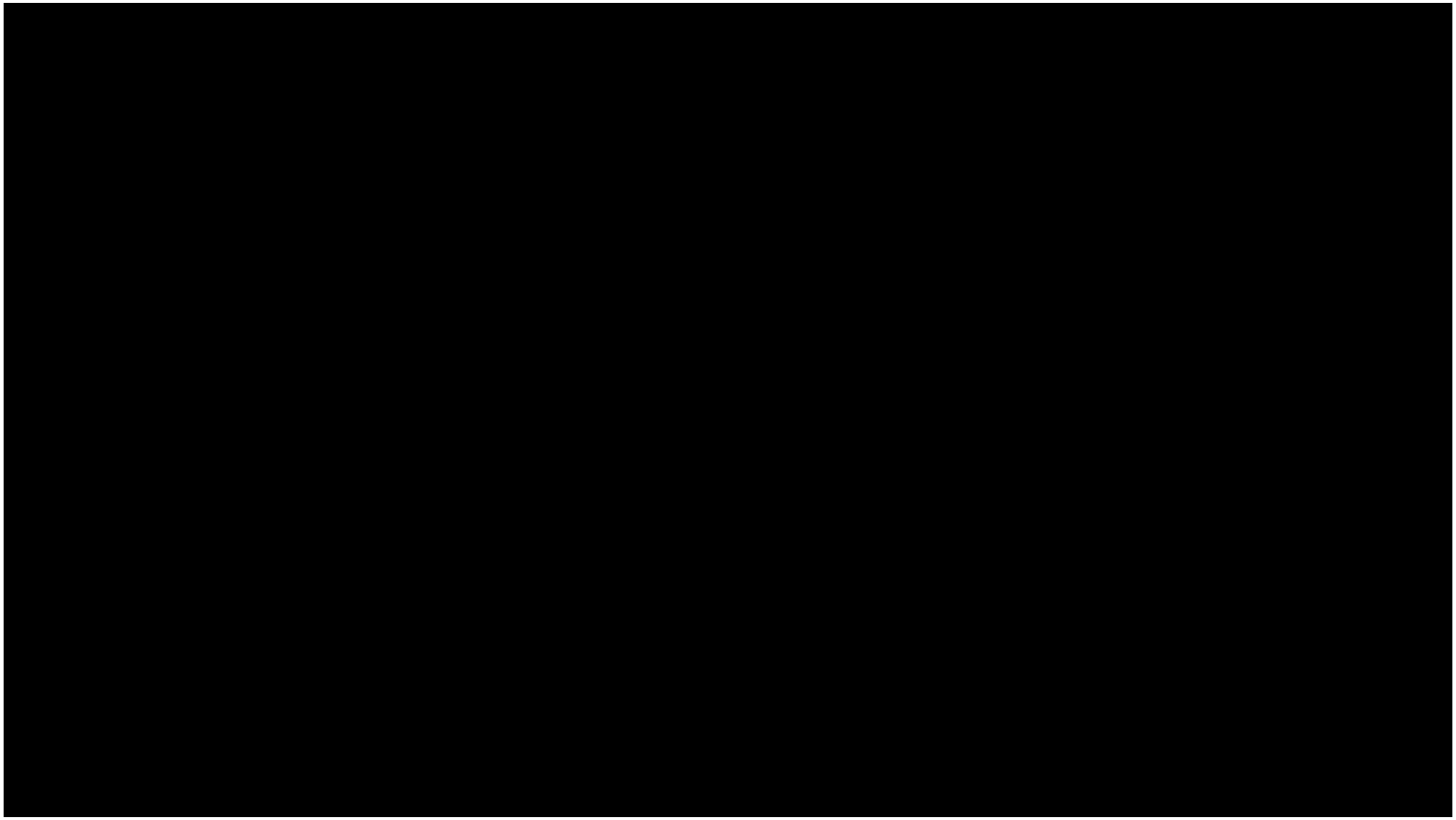
Digital twin in healthcare

Future of healthcare

Future of healthcare

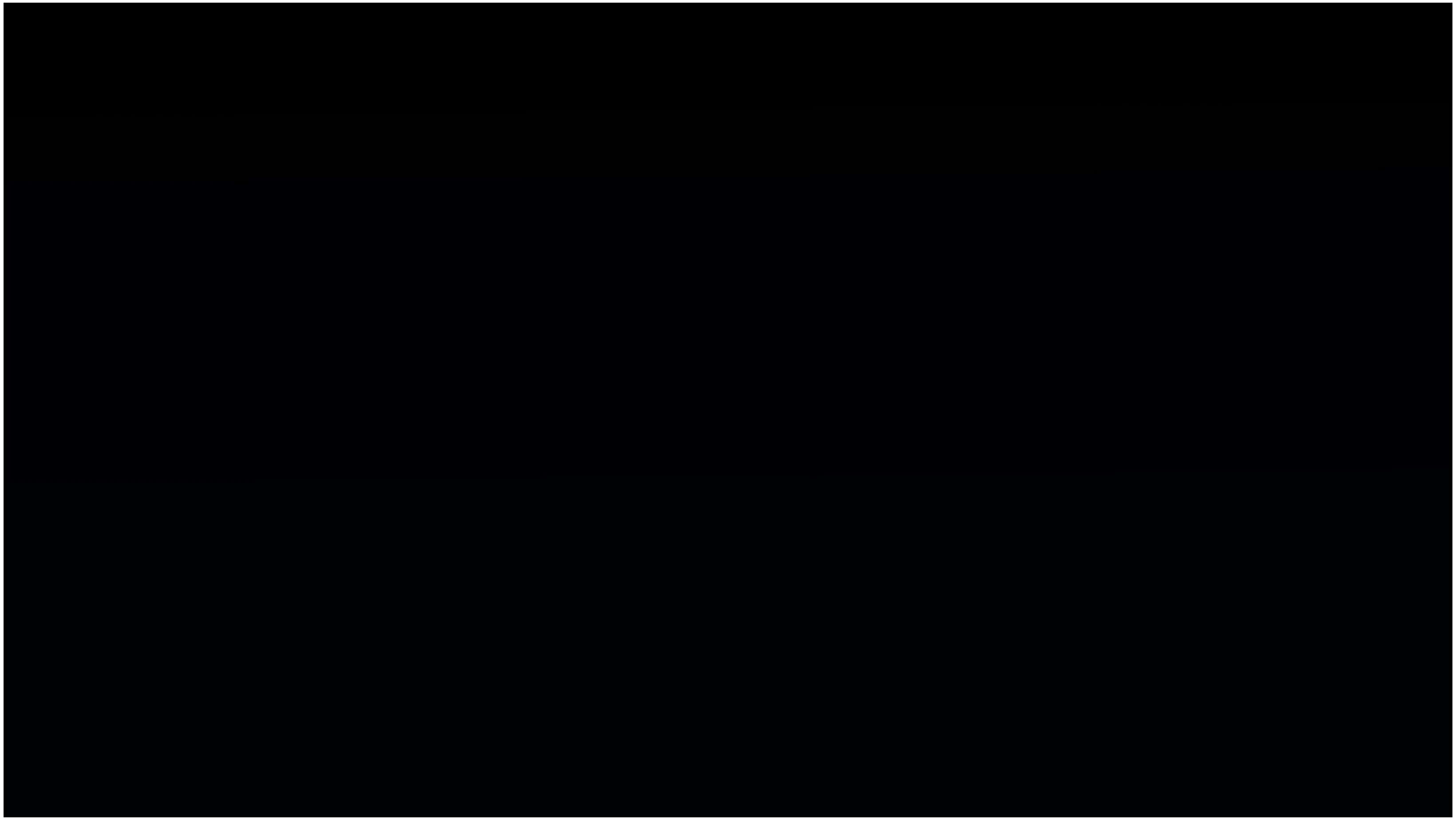
IBM

Future of healthcare: IBM



Deloitte

Future of healthcare: Deloitte



Roche

Future of testing: Roche

Google

Future of healthcare: Google

Google Health

Future of healthcare: Google

Future of healthcare: Google

Google

GE

ediTec

L-Università ta' Malta

Future of healthcare: GE

Siemens Healthineers

Future of healthcare: Siemens

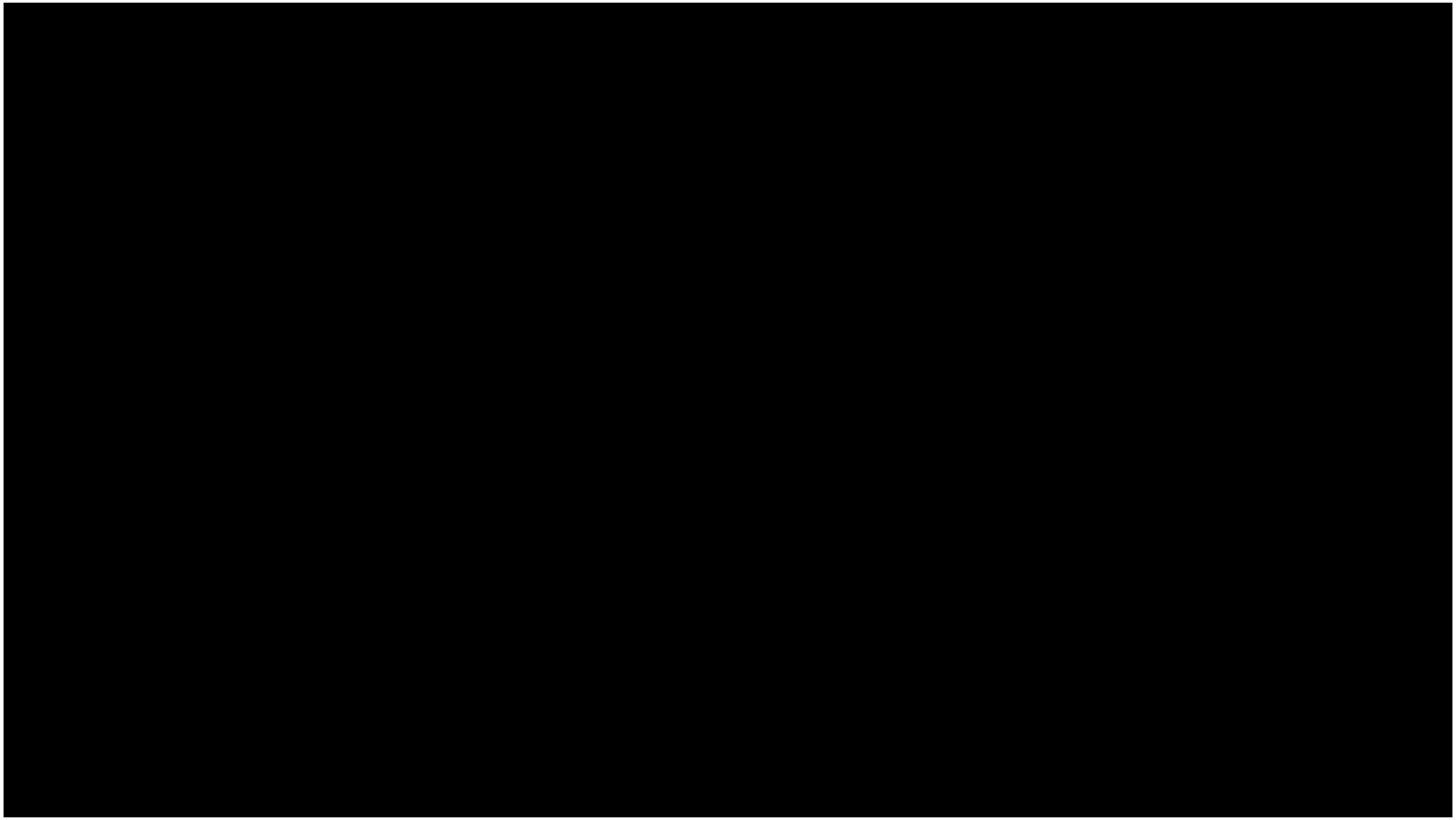
Alcatel-Lucent

Future of healthcare: Alcatel-Lucent

Alcatel·Lucent Enterprise

Philips

Future of healthcare: Philips



Julius Baer

Future of healthcare: Julius Baer

Finland

Future of healthcare: Finland

Anonymity Preserving IoT-Based Contact Tracing Model

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

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[1], (Member, IEEE), EMEKA CHUKWU[1], (Graduate Student Member, IEEE), NIDAL NASSER[2], (Senior Member, IEEE), CHINMAY CHAKRABORTY[3], AND GAURAV GARG[4], (Member, IEEE)

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

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

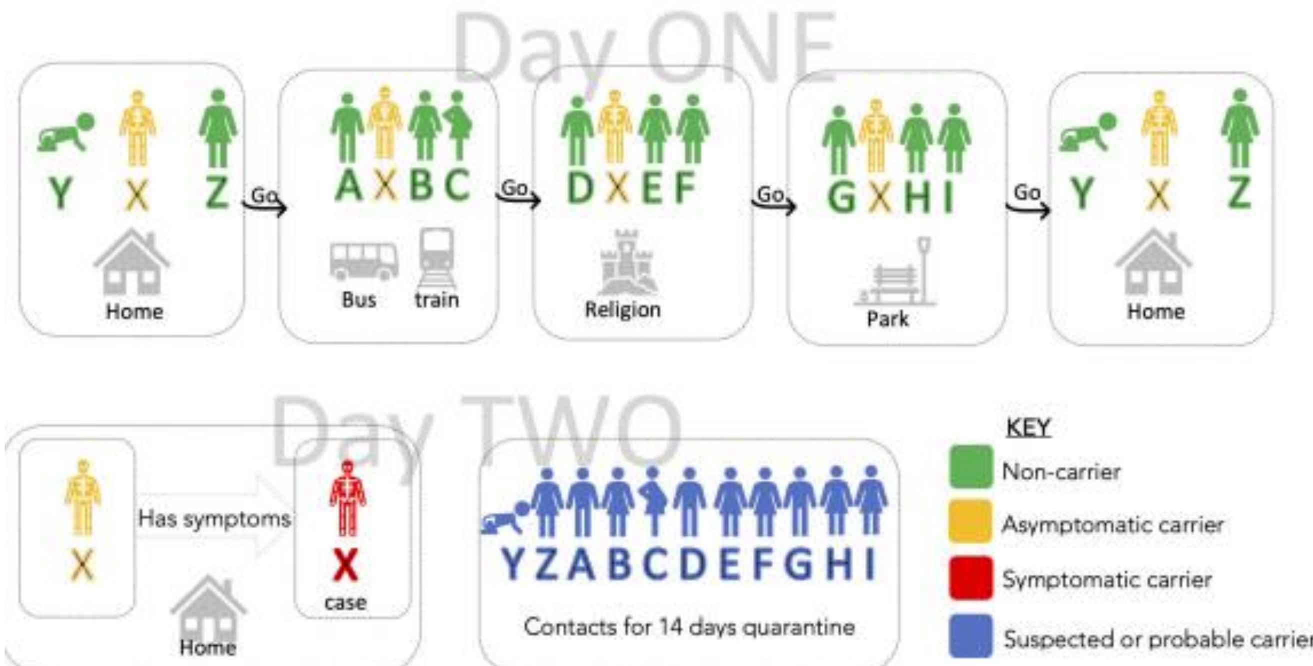
This work was supposed in part by Alfaisal University under Internal Research Grant C20220 and in part by the University of Malta's Research Innovation & Development Trust (RIDT) under Research Grant E18LO77-01.

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, citizen-application, 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

Anonymity Preserving IoT-Based Contact Tracing Model

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



L-Università ta' Malta

Anonymity Preserving IoT-Based Contact Tracing Model

Collaborative partners:

Lalit Garg[1], Emeka Chukwu[1], Nidal Nasser[2], Chinmay Chakraborty[3], Gaurav Garg[4]

[1]University of Malta

[2]Alfaisal University, Riyadh, Saudi Arabia

[3]Birla Institute of Technology, Ranchi, India

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

Anonymity Preserving IoT-Based Contact Tracing Model

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

Anonymity Preserving IoT-Based Contact Tracing Model

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.

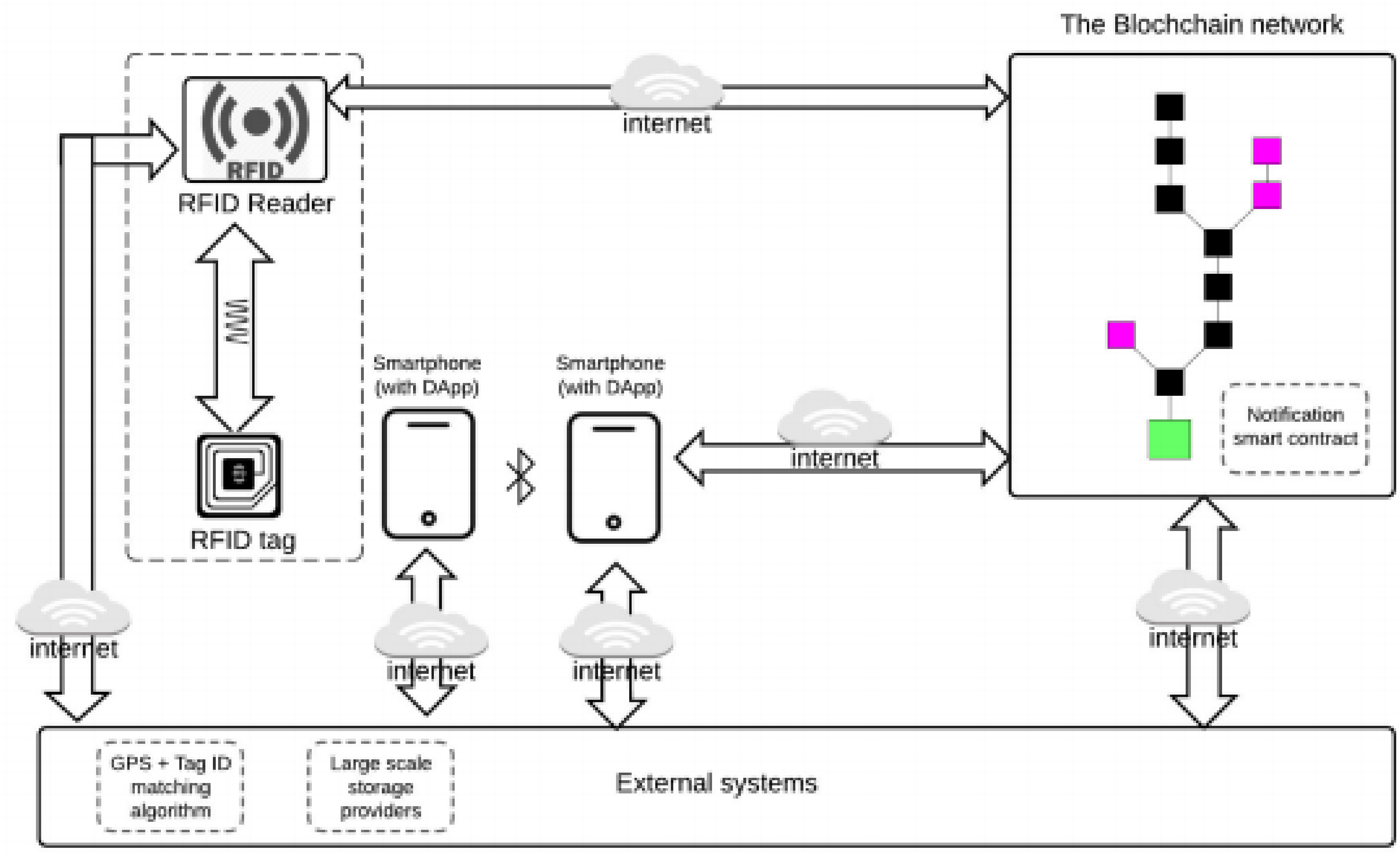
Anonymity Preserving IoT-Based Contact Tracing Model

1. 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

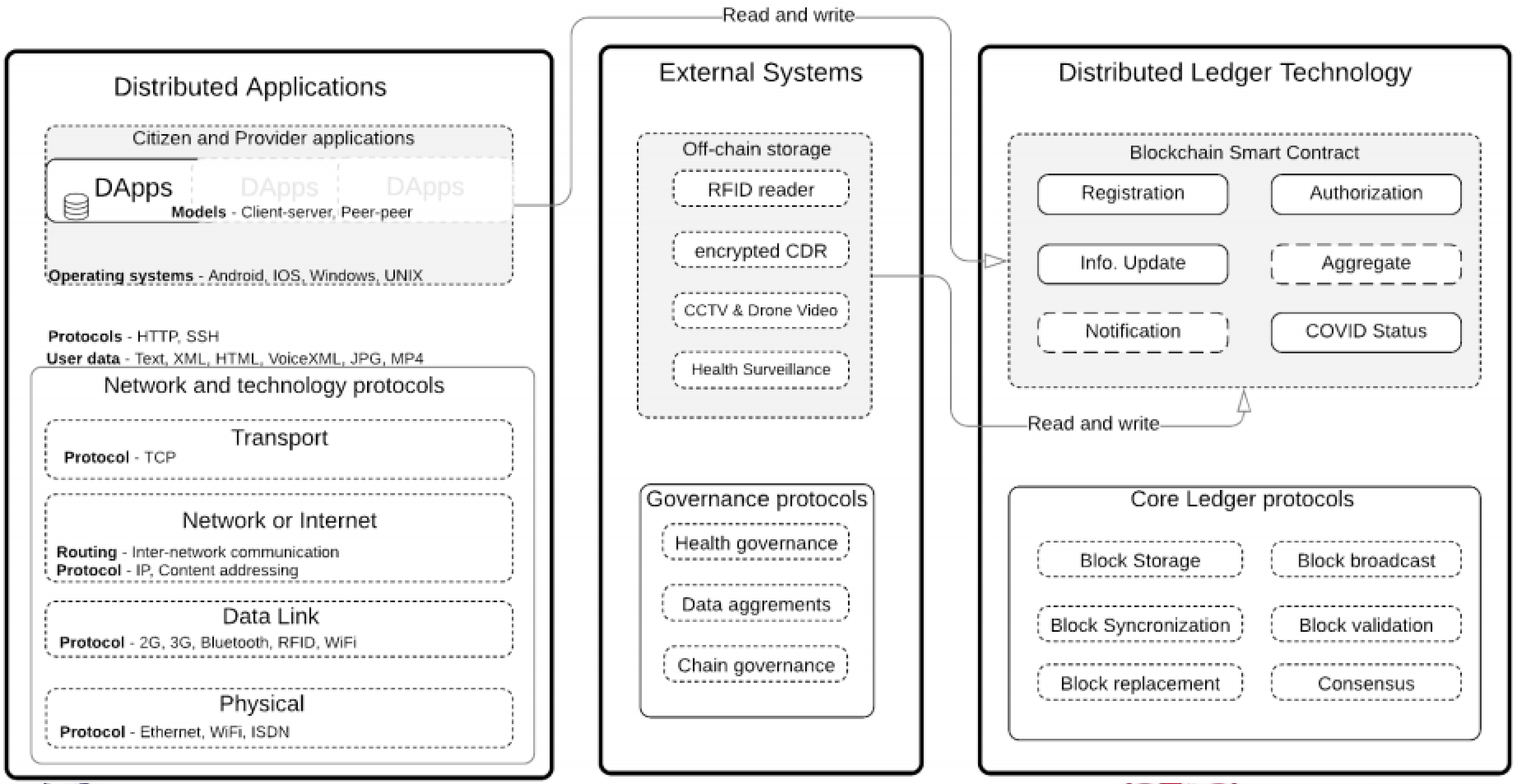
Anonymity Preserving IoT-Based Contact Tracing Model

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.

A masked-image recognition system

A Masked-image recognition system

Lalit Garg,

Emeka Chukwu

ediTec

L-Università 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

Healthcare self-service Kiosk

Lalit Garg,

Emeka Chukwu

L-Università ta' Malta

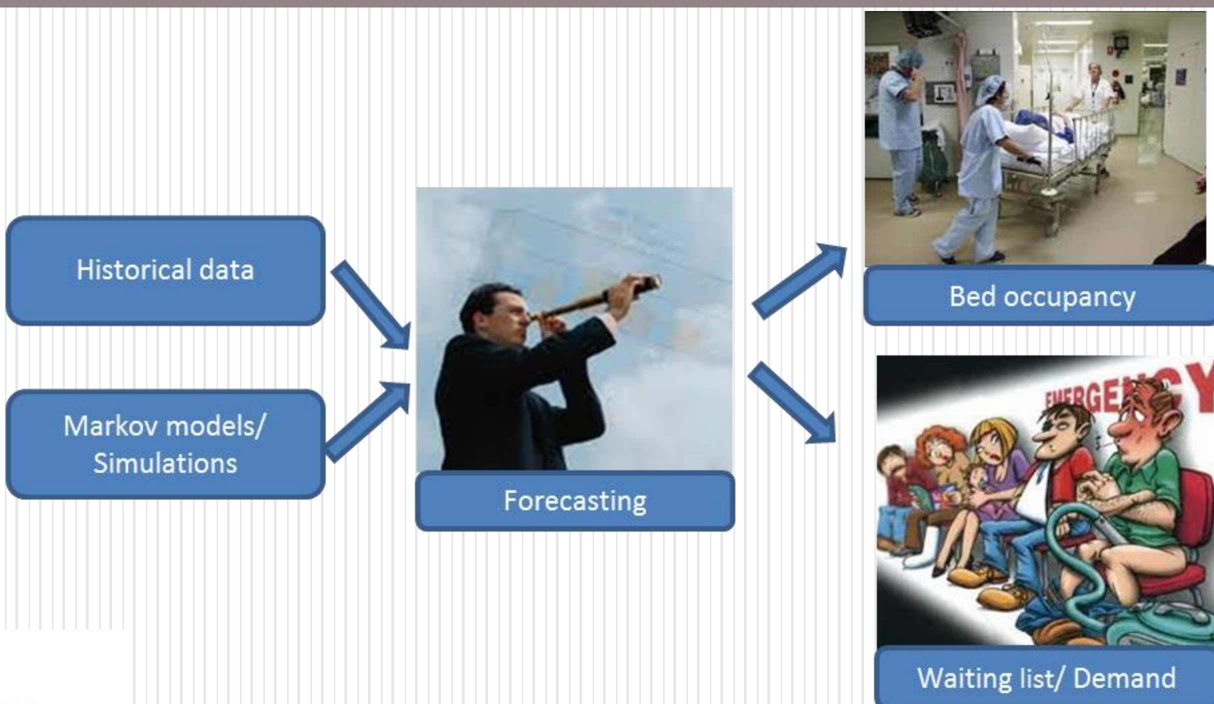
ediTec

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

EFFECTIVE 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,

Health Care Manag Sci
DOI 10.1007/s10729-009-9120-0

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

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 non-homogeneous model is developed which incorporates time-dependent covariates, namely a patient's present age and the present calendar year. 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 optimum utiliza-

INFORMATICA, 2011, Vol. 22, No. 1, 57–72

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

Lalit GARG[1], Sally McCLEAN[1], Brian J. MEENAN[1],
Peter MILLARD[2]

[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 gender, age at the time of admission and primary diagnosis code. Likelihood

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 1SA, 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

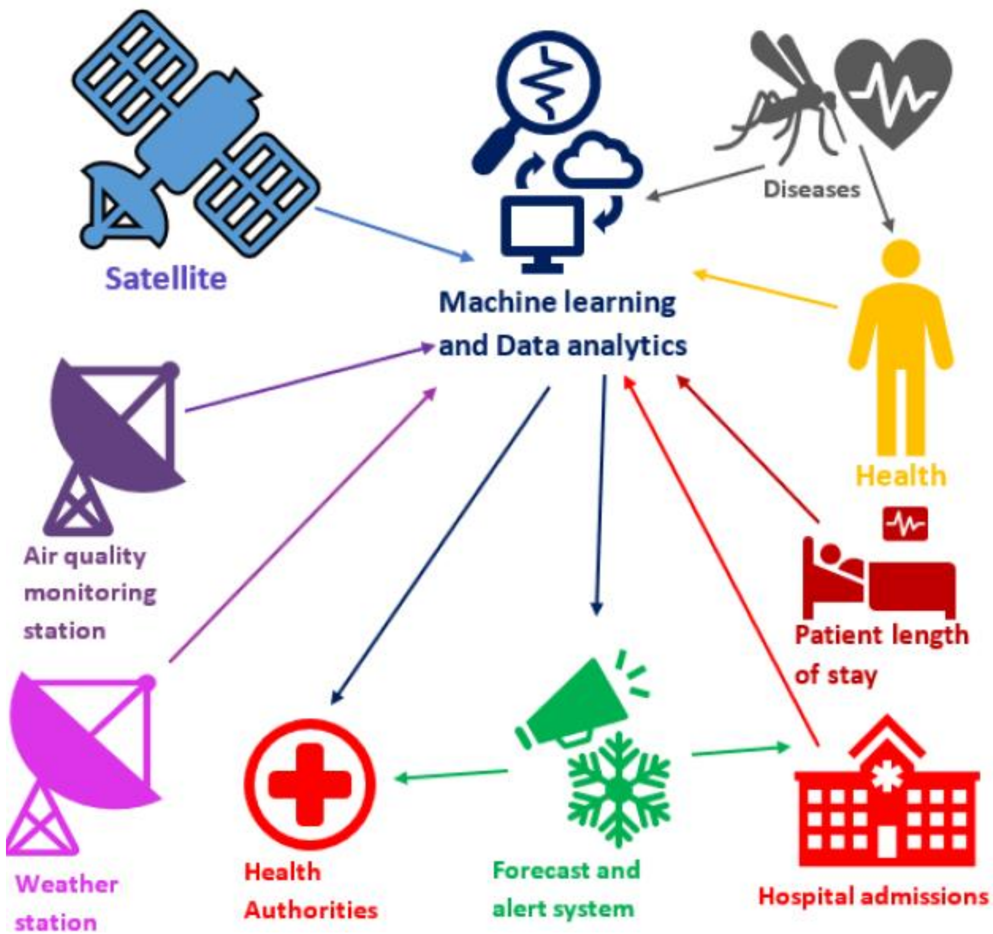
[Received May 2007; accepted May 2008]

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



L-Università ta' Malta

References

- 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.

References

- 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.

References

- 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.

References

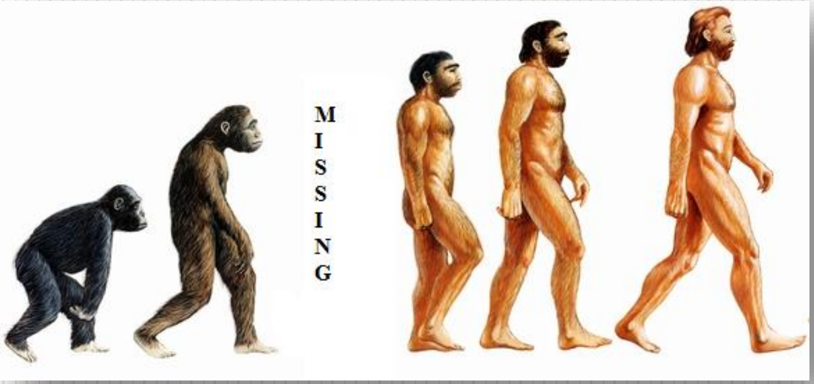
- Barton M, McClean SI, Garg L, Fullerton K (2009). Modelling Stroke Patient Pathways using Survival Analysis and Simulation Modelling. Proceedings of the XIII International Conference on Applied Stochastic Models and Data Analysis (ASMDA 2009), pp. 370-373, Eds: Leonidas Sakalauskas, Christos Skiadas, Edmundas K. Zavadskas, ISBN: 978-9955-28-463-5, Publisher: Vilnius Gediminas Technical University Press. scHolarcitations: 4.

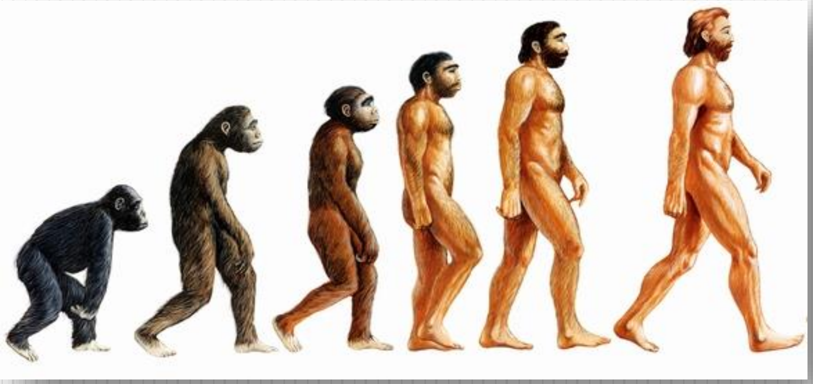
- McClean SI, Garg L, Barton M, Fullerton K, Millard PH (2009). Using Markov Systems to plan Stroke Services. In Intelligent Patient Management (Eds. McClean, S.; Millard, P.; El-Darzi, E.; Nugent, C.D.), Book Series: Studies in Computational Intelligence, Vol. 189. pp. 241-256. Journal H-Index(SJR:1998-2010): 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.

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-

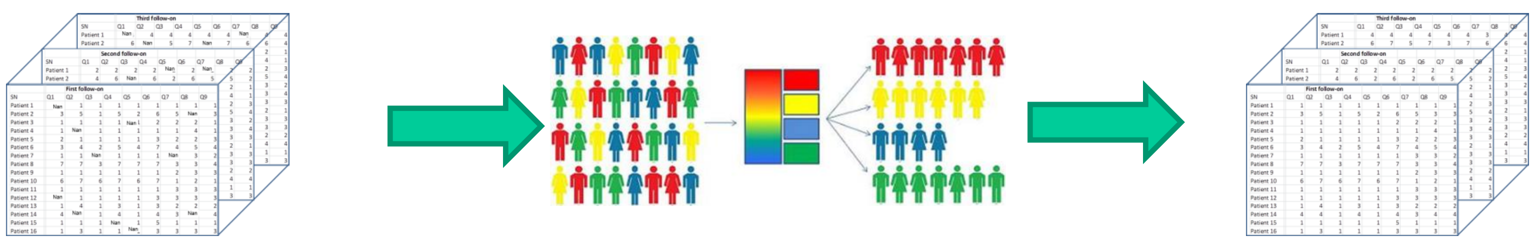
Web-based tools for Missing data handling in medical questionnaires

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

Completed medical questionnaires



4/20/2022

L-Università ta' Malta

Dr Lalit Garg

Web-based tools for Missing data handling in medical questionnaires

Collaborative partners:

Lalit Garg, Justin Dauwels[1], Arul Earnest[2,3], Leong Khai Pang[3]

[1]Nanyang Technological University, Singapore

[2]Duke-NUS Graduate Medical School, Singapore

[3]Tan Tock Seng Hospital (TTSH), Singapore

CP based missing data imputation method

Third follow-on									
SN	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
Patient 1	Nan	4	4	4	4	4	Nan	4	4
Patient 2	6	Nan	5	7	Nan	7	6	6	4
								2	1
								4	1
								2	3
								5	4
								3	2
								3	4
								3	3
								2	1
								3	3
								3	3
								2	2
								4	4
								1	1
								3	3

Second follow-on									
SN	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
Patient 1	2	2	2	2	Nan	2	Nan	2	2
Patient 2	4	6	Nan	6	2	6	5	5	2
								2	1
								4	1
								2	3
								5	4
								3	2
								3	4
								3	3
								2	1
								3	3
								3	3
								2	2
								4	4
								1	1
								3	3

First follow-on									
SN	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
Patient 1	Nan	1	1	1	1	1	1	1	1
Patient 2	3	5	1	5	2	6	5	Nan	3
Patient 3	1	1	1	1	Nan	2	2	2	1
Patient 4	1	Nan	1	1	1	1	1	4	1
Patient 5	2	1	1	1	1	3	2	2	3
Patient 6	3	4	2	5	4	7	4	5	4
Patient 7	1	1	Nan	1	1	1	Nan	3	2
Patient 8	7	7	3	7	7	7	3	3	4
Patient 9	1	1	1	1	1	1	2	3	3
Patient 10	6	7	6	7	6	7	1	2	1
Patient 11	1	1	1	1	1	1	3	3	3
Patient 12	Nan	1	1	1	1	3	3	3	3
Patient 13	1	4	1	3	1	3	2	2	2
Patient 14	4	Nan	1	4	1	4	3	Nan	4
Patient 15	1	1	1	Nan	1	5	1	1	1
Patient 16	1	3	1	1	Nan	3	3	3	3

CP based missing data imputation method

Third follow-on

SN	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
Patient 1	4	4	4	4	4	4	3	4	4
Patient 2	6	7	5	7	3	7	6	6	4
								2	1
								4	1
								2	3
								5	4
								3	2
								3	4
								3	3
								2	1
								3	3
								3	3
								2	2
								4	4
								1	1
								3	3

Second follow-on

SN	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
Patient 1	2	2	2	2	2	2	2	2	2
Patient 2	4	6	2	6	2	6	5	5	2
								2	1
								4	1
								2	3
								5	4
								3	2
								3	4
								3	3
								2	1
								3	3
								3	3
								2	2
								4	4
								1	1
								3	3

First follow-on

SN	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
Patient 1	1	1	1	1	1	1	1	1	1
Patient 2	3	5	1	5	2	6	5	3	3
Patient 3	1	1	1	1	1	2	2	2	1
Patient 4	1	1	1	1	1	1	1	4	1
Patient 5	2	1	1	1	1	3	2	2	3
Patient 6	3	4	2	5	4	7	4	5	4
Patient 7	1	1	1	1	1	1	3	3	2
Patient 8	7	7	3	7	7	7	3	3	4
Patient 9	1	1	1	1	1	1	2	3	3
Patient 10	6	7	6	7	6	7	1	2	1
Patient 11	1	1	1	1	1	1	3	3	3
Patient 12	1	1	1	1	1	3	3	3	3
Patient 13	1	4	1	3	1	3	2	2	2
Patient 14	4	4	1	4	1	4	3	4	4
Patient 15	1	1	1	1	1	5	1	1	1
Patient 16	1	3	1	1	1	3	3	3	3

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&isnumber=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

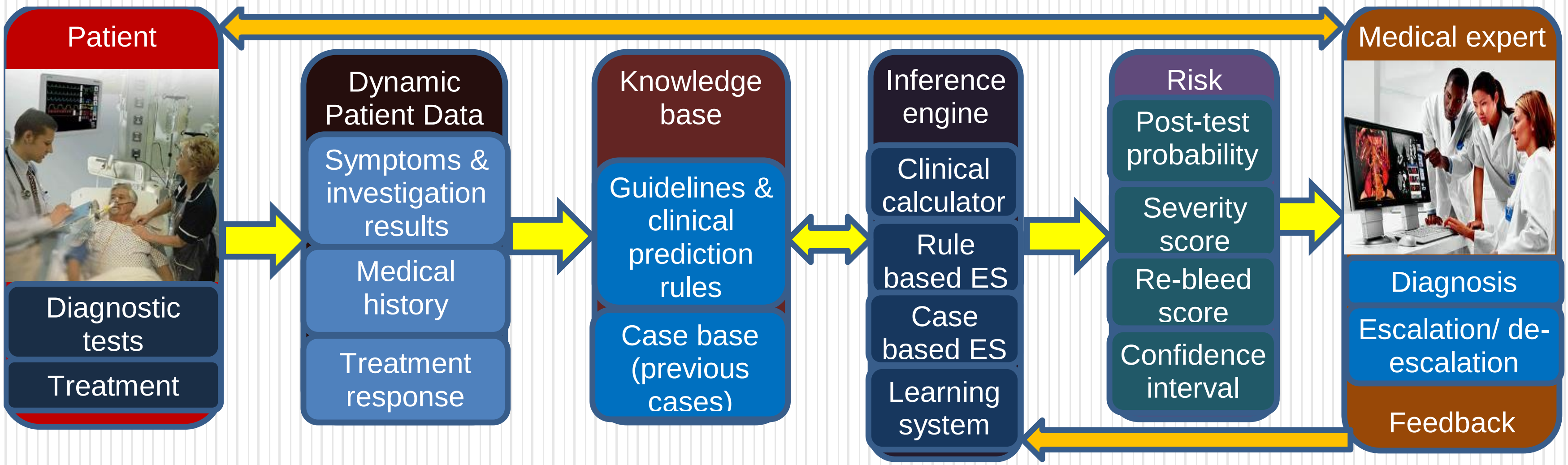
More info…

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



ediTec

L-Università ta' Malta

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 Garg[1], Giovanni Masala[2], Sally I. McClean[1], Marco Micocci[2], Giuseppina Cannas[2]
[1]*School of Computing and Information Engineering, University of Ulster, Coleraine, UK*
[2]*Faculty 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 Artificial Health Intelligence Projects

Smart Sensor for EEG Acquisition and Epileptic Seizure Detection and prediction

Spectral Energy Calculation (SVM) Or Empirical mode decomposition (EMD) → Feature Extraction SVM/ELM / probabilistic clustering → Seizure detection and prediction

ediTec

L-Università ta' Malta

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, Msida, 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

M. Pant 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

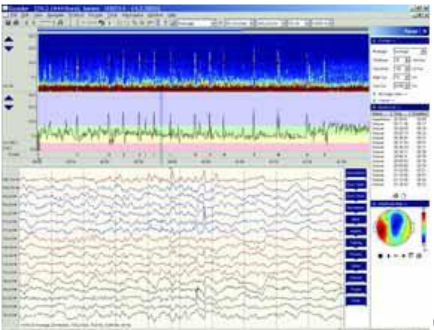
Data: Massachusetts General Hospital, MIT, USA.

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.

EEG and fMRI integration based models of brain disorders

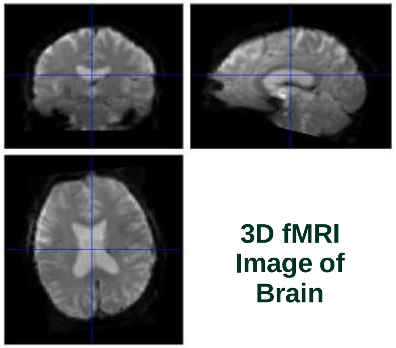


Electroencephalogram (EEG)

Pros: Fast Temporal Response

Cons: Poor Spatial Resolution

(CPP and 2-D)



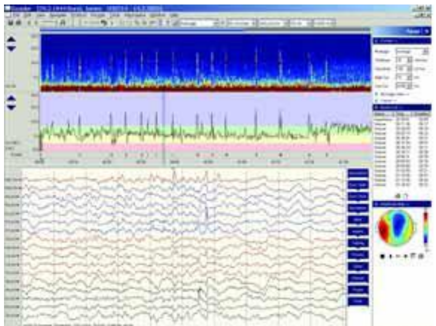
3D fMRI Image of Brain

functional Magnetic Resonance Imaging(fMRI)

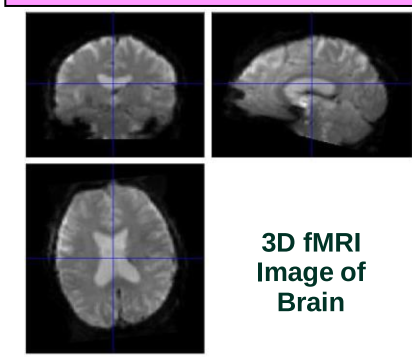
Pros: Good Spatial Resolution (3D)

Cons: Slow BOLD transient response

EEG and fMRI integration based models of brain disorders



Combined EEG-fMRI Analysis



3D fMRI Image of Brain

Localization for various signals on BOLD Activity Brain Maps (corresponding to the EEG signals from individual channels)

Defining the Model for Various Mental Disorders and Degenerations

Analyse and Modify for various Mental Disorders and Degenerations

EEG and fMRI integration based models of brain disorders

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)

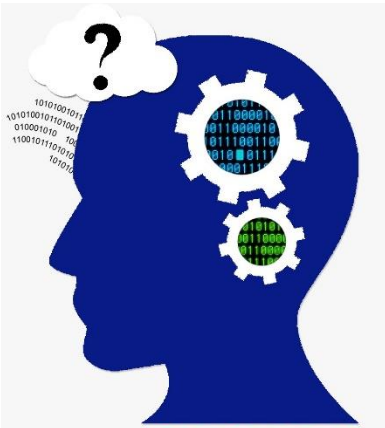
EEG and fMRI integration based models of brain disorders

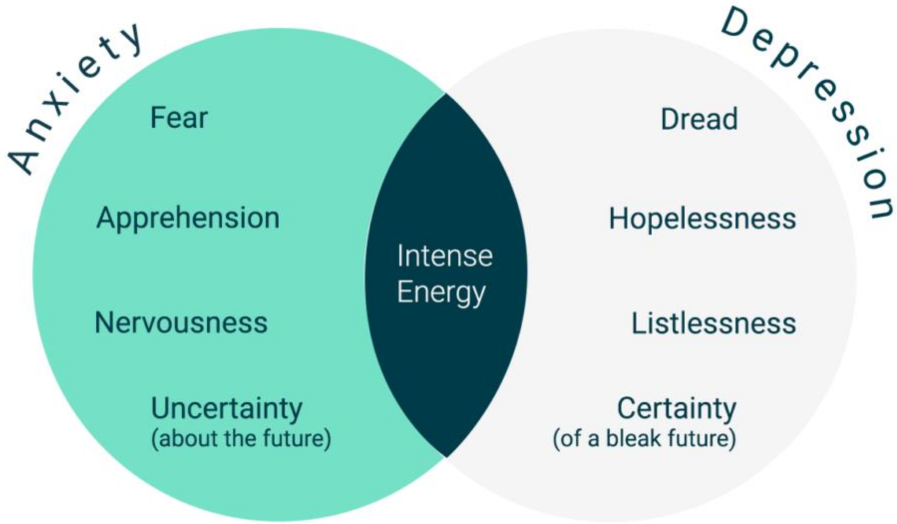
More info:

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.*

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

Predicting Neurological Disorder via Social Media





ediTec

L-Università ta' Malta

Dr Lalit Garg

Having anxiety and depression...

Depression: Just lay in bed all day and do nothing. You're life is worthless anyway.

Me: Okay.

Anxiety: What the hell are you doing? You need to study or you'll fail all your classes, drop out of school, and end up living on the street with no friends!!

Depression: No stay here with me.

Me: ?????

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)

ediTec

L-Università
ta' Malta

Dr Lalit Garg