Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 154 (2019) 361–368

Procedia Computer Science

www.elsevier.com/locate/procedia

8th International Congress of Information and Communication Technology, ICICT 2019

Machine Learning Approaches in Smart Health

Zeina Rayan [a] *, Marco Alfonse [a], Abdel-Badeeh M. Salem [a]

[a] *Department of Computer Science, Faculty of computer and information science, University of Ain Shams, Cairo, Egypt*

Abstract

The increase of age average led to an increase in the demand of providing and improving the service of healthcare. The advancing of the information and communication technology (ICT) led to the development of smart cities which have a lot of components. One of those components is Smart Health (s-Health), which is used in improving healthcare by providing many services such as patient monitoring, early diagnosis of diseases and so on. Nowadays there are many machine learning techniques that can facilitates s-Health services. This paper reviews recent published papers in the area of smart health starting from the years 2011 to 2017, and a structured analysis for different machine learning (ML) approaches that are applied in s-Health. The results show that the ML approach is used in many s-Health applications such as Glaucoma diagnosis, Alzheimer's disease, bacterial sepsis diagnoses, the Intensive Care Unit (ICU) readmissions, and cataract detection. The Artificial Neural Network (ANN), Support Vector Machine (SVM) algorithm and deep learning models especially the Convolutional Neural Network (CNN) are the most commonly used machine learning approaches where they proved to get high evaluation performance in most cases.

© 2019 The Authors. Published by Elsevier Ltd.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/)
Selection and peer-review under responsibility of the 8th International Congress of Information and Communication Technology, ICICT 2019.

Keywords: Smart Health; Electronic Health; Medical Informatics; Machine Learning

1. Introduction

Smart Health is a field which grown as a subset of two fields, smart cities and Electronic Health (e-Health) 1. Smart cities can be defined as "cities strongly founded on information and communication technologies that invest in human and social capital to improve the quality of life of their citizens by fostering economic growth,

* Corresponding author : Zeina Rayan
*E-mail address:*zeinarayan@cis.asu.edu.eg

1877-0509 © 2019 The Authors. Published by Elsevier Ltd.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/)
Selection and peer-review under responsibility of the 8th International Congress of Information and Communication Technology, ICICT 2019.
10.1016/j.procs.2019.06.052

participatory governance, wise management of resources, sustainability, and efficient mobility, whilst they guarantee the privacy and security of the citizens."2. e-Health can be defined as "an emerging field in the intersection of medical informatics, public health and business, referring to health services and information delivered or enhanced through the Internet and related technologies. In a broader sense, the term characterizes not only a technical development, but also a state-of-mind, a way of thinking, an attitude, and a commitment for networked, global thinking, to improve health care locally, regionally, and worldwide by using information and communication technology."3. There is an intersection between s-Health and Mobile Health (m-Health); m-Health can be defined as "emerging mobile communications and network technologies for healthcare systems"4.

Machine Learning (ML) is a field that grew out of Artificial Intelligence (AI). It is concerned with designing and developing algorithms that enable the computers to evolve their behaviors according to empirical data. The ML approach is evolving rapidly as a result of the improvement of the ML algorithms, enhanced methods of capturing data, improved computer networks, new sensors/IO units, and the interest in self-customization to users' behavior 5. ML plays an important role in s-Health, it can improve the quality of healthcare services by providing accurate medical diagnosis, predicting diseases in early stages and disease analyses.

This paper is structured as follows: section 2 presents the components of a smart health system, and the pipeline of developing it, section 3 contains a comparative study between different machine learning approaches that are applied in s-Health, section 4 presents the conclusions and future work.

2. Smart Health

s-Health is a new form of healthcare which is a subfield of e-Health using Electronic Health Records (EHR) and other variables coming from the smart city's infrastructure; in order to improve the healthcare. The idea of a smart health system is that it uses all data coming from sensors on the patient body, smart homes, smart city infrastructure, and robots to help make better decisions and improve healthcare by providing emergency response and paging doctors, nurses, and technicians (see Fig. 1). It can also support self-diagnosis, monitoring, early detection and treatments.

Fig. 1 The components of a smart health system

The smart health system can be described as a pipeline that includes data acquisition, networking and computing technologies, data security and privacy, data processing, and data dissemination. Fig. 2 shows the pipeline for developing a smart heath system.

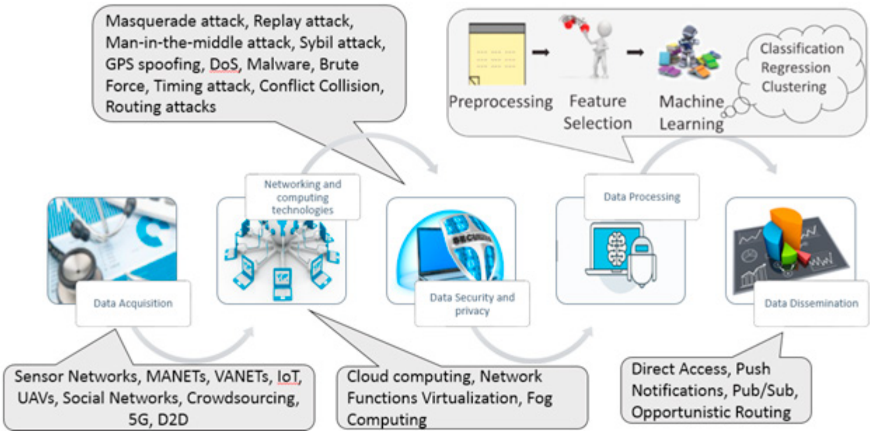

Fig. 2 Smart health pipeline

The data acquisition means collecting data from different sources such as sensor network 6,7,8, Mobile Ad hoc Networks 9,10 , Vehicular Ad hoc Networks 11,12, Social networks 13, Internet of Things (IoT) 14, 5G devices 15,16, Device-to-Device (D2D) 17, Unmanned Aerial Vehicles (UAVs) 18, or may be a combination of a set of them. The networking and computing technologies are applied to the data gathered in the data acquisition phase, since this data is a bit complex, we need a server or a computing technology where the data can be processed such as cloud computing 19, or fog computing 20. The data security and privacy plays a big role in both s-Health and smart cities where gathering so many information about citizens could possibly violate the citizen privacy 21. The data processing involves preprocessing to remove the noise and clean the data, feature extraction to find the relevant features, and ML technique to process the data and perform the required task such as classification/clustering …etc. The data dissemination is responsible for providing the output of the data processing phase to the target parties by means of direct access, push notifications, pub/sub, or opportunistic routing 22.

3. ML Approaches in Smart Health

This section presents a comparative study between different machine learning approaches that are applied in s-Health applications and systems.

Chai, et. al [23] proposed a deep-learning model using CNN for Glaucoma diagnosis, using data which was collected from Beijing Tongren Hospital in China. The dataset includes 3554 images of retinal fundus collected from 2000 patients who suffered from numerus types of eye diseases. This data set contains 1391 images that are diagnosed with glaucoma, and the remaining images (2163) are not. The proposed model gives accuracy 81.69%.

Zhang, et. al [24] proposed a model to predict the Mini-Mental State Examination (MMSE) scores in Alzheimer’s disease (AD) based on multi-granularity whole-brain segmentation volumes through images. They used the Alzheimer's disease Neuroimaging Initiative (ADNI) database, which includes healthy control cases (190), mild cognitive impairment cases (331) and AD cases (157). The proposed model gives Area Under Curve (AUC) 0.554.

Liu, Y. and Choi, K.S. [25] proposed a model to diagnose the bacterial sepsis in critically ill patients. The data was collected from Medical ICU (MICU) and ICU departments of the General Hospital of Guangzhou, China (185 inpatients). The proposed model gives accuracy 90.8%.

Viegas, R., et. al [26] proposed an ensemble model to predict readmissions of ICU utilizing feature selection and fuzzy modeling approaches. They used the Multi–Parameter Intelligent Monitoring for Intensive Care (MIMIC II) database, which is an archive of data, publically available, collected from more than 32,000 ICU patients within 2001 and 2008 [27]. The proposed model gives AUC 0.79.

Dong, Y., et. al [28] proposed a model to use the information of retinal vascular for the detection of cataract, which is based on the retinal images classification. The dataset used contains 445 fundus images (199 normal images, 148 slight images, 71 images are medium and 27 images are severe). The classification of these fundus images are performed by the professional ophthalmologist, the data is divided into 2 sets (a set contains 70% of the data that are randomly selected are used for training the classification algorithm, and the remaining 30% of the data is used to test the model performance). The proposed model gives accuracy 94.91%.

Zheng, B., et. al [29] proposed a model to study the hospital readmissions risk prediction. The dataset was collected from medical records of different hospitals, which contains 1641 records (316 records of them are readmitted to hospitals during 30 days of discharge). The proposed model gives accuracy 83.8%.

Fialho, A.S., et. al [30] proposed a model to predict the readmission of ICU between 24 and 72 hours after the discharge from ICU, they used the MIMIC II database. The proposed model gives accuracy 74%.

Table 1 shows a comparison between different machine learning approaches used in smart health.

Table 1 ML approaches in s-Health

Authors	Objective	Data Set	Preprocessing	Feature Extraction	Machine Learning Technique	Evaluation
Chai, et. al, 2017 [23]	Glaucoma Diagnosis	Data collected from Beijing Tongren Hospital.	Equalize the size of images (Normalization)	Scale-Invariant Feature Transform (SIFT) features	Logistic regression	Accuracy = 60.84%
				Local Binary Patterns (LBP) features	Logistic regression	Accuracy = 62.67%
			Normalization and Segmentation	whole image	One-branch CNN (5-conv layers)	Accuracy = 71.36%
				extracted image	One-branch CNN (5-conv layers)	Accuracy = 74.95%
					Two-branch CNN (six-conv layers)	Accuracy = 74.89%
					Two-branch CNN (5-conv layers)	Accuracy = 81.69%
Zhang, et. al, 2017 [24]	Prediction of the MMSE scores in Alzheimer's disease.	ADNI database.	segmented each image into 5-level whole-brain segmentations utilizing the Multi-Atlas Likelihood Fusion algorithm	Level 1	Linear Regression	AUC = 0.318
					Ridge Regression	AUC = 0.530
					Lasso Regression	AUC = 0.520
					Elastic-net	AUC = 0.527
				Level 2	Linear Regression	AUC = 0.453
					Ridge Regression	AUC = 0.542
					Lasso Regression	AUC = 0.550
					Elastic-net	AUC = 0.554

				Level 3	Linear Regression	AUC = 0.388
					Ridge Regression	AUC = 0.427
					Lasso Regression	AUC = 0.419
					Elastic-net	AUC = 0.424
				Level 4	Linear Regression	AUC = 0.309
					Ridge Regression	AUC = 0.318
					Lasso Regression	AUC = 0.324
					Elastic-net	AUC = 0.322
				Level 5	Linear Regression	AUC = 0.266
					Ridge Regression	AUC = 0.273
					Lasso Regression	AUC = 0.262
					Elastic-net	AUC = 0.268
Liu, Y. and Choi, K.S., 2017 [25]	Diagnose bacterial sepsis.	Data collected from the General Hospital of Guangzhou, China.	Normalization for the age attribute		Procalcitonin (PCT) (Cutoff = 3.40)	Accuracy = 78.4%
					PCT (Cutoff = 8.56)	Accuracy = 82.2%
					SVM	Accuracy = 88.6%
					Random Forest (RF)	Accuracy = 89.2%
					ANN	Accuracy = 90.8%
Viegas, R., et. al, 2017 [26]	Prediction of ICU readmissions.	MIMIC II	exactly define the scope of allowable qualities as indicated by the expert	Area under the sensitivity and specificity curves	Average ensemble decision criteria	AUC = 0.77±0.02
					Maximum distance ensemble decision criteria	AUC = 0.77±0.02
					Weighted distance ensemble decision criteria	AUC = 0.77±0.02

				Sensitivity and specificity at the intersection threshold	Average ensemble decision criteria	AUC = 0.76±0.02
					Maximum distance ensemble decision criteria	AUC = 0.75±0.02
					Weighted distance ensemble decision criteria	AUC = 0.76±0.02
				Sensitivity and specificity close to the intersection threshold	Average ensemble decision criteria	AUC = 0.75±0.02
					Maximum distance ensemble decision criteria	AUC = 0.74±0.02
					Weighted distance ensemble decision criteria	AUC = 0.75±0.02
Dong, Y., et. al, 2016 [28]	cataract detection	Data collected from different sources		Wavelet features	SVM	Accuracy = 81.81%
				Texture features		Accuracy = 80.33%
					staking algorithm	Accuracy = 84%
Zheng, B., et. al, 2015 [29]	The risk prediction of hospital readmissions.	Data was collected from different hospitals.	Using a range defined for admissible values.		Radial Basis Function Neural Network	Accuracy = 56.1%
					Particle Swarm Optimization–SVM with Radial Basis Function	Accuracy = 83.8%
					Linear SVM	Accuracy = 50.6%
					Polynomial SVM	Accuracy = 52.7%
					RBF-based SVM	Accuracy = 69.5%

Fialho, A.S., et. al, 2012 [30]	Predict ICU readmission.	MIMIC II	• Data missing for an intentional reason was deleted. • Data missing for unintentional reason was given the last available value was used.	sequential forward selection	Data mining approach	Accuracy = 71 ± 3%
				Intellegent tool titeled: Acute Physiology And Chronic Health Evaluation (APACHE) version II		Accuracy = 61 ± 2%
				APACHE version III Modified tool		Accuracy = 65 ± 2%

From the above table, we can find that the ML approach is used in many s-Health applications such as Glaucoma diagnosis, Alzheimer's disease, bacterial sepsis diagnoses, ICU readmissions, and cataract detection. The ANN, SVM algorithm and deep learning models especially CNN are the most commonly used machine learning approaches where they proved to get high evaluation performance in most cases.

4. Conclusions and Future Work

Smart health is a developing and exceedingly critical research field with a possibly noteworthy effect on the conventional healthcare industry. This work presents an overview of the challenges, pipeline, and techniques of smart health. A systematic pipeline of data processing is accommodated for conventional smart health, covering data acquisition, data processing, data dissemination, data security and privacy, and networking and computing technologies. In spite of numerous chances and methodologies for data analytics in healthcare presented in this work, there are numerous different bearings to be investigated concerning different aspects of healthcare data such as quality, privacy and so on. In future work, we are going to apply Machine Learning algorithms to detect anomaly over intensive care patient data.

References

1. Solanas, A., Patsakis, C., Conti, M., Vlachos, I.S., Ramos, V., Falcone, F., Postolache, O., Pérez-Martínez, P.A., Di Pietro, R., Perrea, D.N. and Martinez-Balleste, A., Smart health: a context-aware health paradigm within smart cities. IEEE Communications Magazine, 52(8), pp.74-81, 2014.
2. P. A. Pérez-Martínez, A. Martínez-Ballesté, and A. Solanas, "Privacy in Smart Cities — A Case Study of Smart Public Parking," Proc. 3rd Int'l Conf. Pervasive Embedded Computing and Commun. Sys., pp. 55-59, 2013.
3. G. Eysenbach, "What Is e-Health?" J. Medical Internet Research, vol. 3, no. 2, pp. 20-21, Apr-June 2001.
4. Istepanian, Robert, Swamy Laxminarayan, and Constantinos S. Pattichis. M-health. New York, NY: Springer Science+ Business Media, Incorporated, 2006.
5. John D. Kelleher, Brian Mac Namee, Aoife D'Arcy, Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies, 1st Edition, The MIT Press, 2015.
6. Y. Fujii, N. Yoshiura, A. Takita, and N. Ohta, "Smart street light system with energy saving function based on the sensor network," in Proceedings of the Fourth International Conference on Future Energy Systems, ser. e-Energy '13, New York, NY, USA, pp. 271–272, 2013. [Online]. Available: http://doi.acm.org/10.1145/2487166.2487202
7. N. Yoshiura, Y. Fujii, and N. Ohta, "Smart street light system looking like usual street lights based on sensor networks," in 13th IEEE International Symposium on Communications and Information Technologies, pp. 633–637, 2013.
8. Abinaya, R., V. Varsha, and Kaluvan Hariharan. "An intelligent street light system based on piezoelectric sensor networks." In 4th International Conference on Electronics and Communication Systems (ICECS), pp. 138-142. IEEE, 2017.

A. Didolkar and F. Zama, "SMARTIE: A security solution for a smart city using internet of things - architecture reference model," International Journal of Advance Research and Innovative Ideas in Education, vol. 2, no. 4, pp. 231-236, 2016.

9. W. A. Jabbar, M. Ismail, and R. Nordin, “On the performance of the current MANET routing protocols for VoIP, HTTP, and FTP applications,” Journal of Computer Networks and Communications, Volume 2014, Article ID 154983, 2014.
10. S. Al-Sultan, M. M. Al-Doori, A. H. Al-Bayatti, and H. Zedan, “A comprehensive survey on vehicular ad hoc network,” Journal of Network and Computer Applications, vol. 37, pp. 380-392, 2014.
11. K. Mershad and H. Artail, “A framework for secure and efficient data acquisition in vehicular ad hoc networks,” IEEE Transactions on Vehicular Technology, vol. 62, no. 2, pp. 536-551, Feb 2013.
12. E. Psomakelis, F. Aisopos, A. Litke, K. Tserpes, M. Kardara, and P. M. Campo, “Big IoT and social networking data for smart cities: Algorithmic improvements on big data analysis in the context of radical city applications,” arXiv preprint arXiv: 1607.00509, 2016.
13. Kumar, S.P., Samson, V.R.R., Sai, U.B., Rao, P.M. and Eswar, K.K. Smart health monitoring system of patient through IoT. In International Conference on IoT in Social, Mobile, Analytics and Cloud (I-SMAC), pp. 551-556, IEEE, 2017.
    A. Orsino, G. Araniti, L. Militano, J. Alonso-Zarate, A. Molinaro, and A. Iera, “Energy efficient IoT data collection in smart cities exploiting D2D communications,” Sensors, vol. 16, no. 6, p. 836-851, 2016.
14. Beard Cory, "High Reliability 4G and 5G Cellular Wireless Services for Smart Cities", IEEE Smart Cities Kansas City Kickoff Workshop, February 2016.
15. M. N. Tehrani, M. Uysal, and H. Yanikomeroglu, “Device-to-Device communication in 5g cellular networks: challenges, solutions, and future directions,” IEEE Communications Magazine, vol. 52, no. 5, pp. 86–92, 2014.
16. G. Ermacora, S. Rosa, and A. Toma, “Fly4smartcity: A cloud robotics service for smart city applications,” Journal of Ambient Intelligence and Smart Environments, vol. 8, no. 3, pp. 347–358, 2016.
17. Mell, Peter, and Tim Grance. "The NIST definition of cloud computing", Computer Security Division, Information Technology Laboratory, the National Institute of Standards and Technology, Gaithersburg, 2011.
18. Bonomi, F., Milito, R., Zhu, J. and Addepalli, S., August. Fog computing and its role in the internet of things. In Proceedings of the first edition of the MCC workshop on Mobile cloud computing, ACM, pp. 13-16, 2012.
    A. Martínez-Ballesté, P. A. Pérez-Martínez, and A. Solanas, “The Pursuit of Citizens’ Privacy: A Privacy-Aware Smart City is Possible,” IEEE Communication Magazine, vol. 51, no. 6, pp. 136-141, 2013.
19. Gharaibeh, A., Salahuddin, M.A., Hussini, S.J., Khreishah, A., Khalil, I., Guizani, M. and Al-Fuqaha, A. Smart Cities: A Survey on Data Management, Security, and Enabling Technologies. IEEE Communications Surveys and Tutorials, 19(4), pp.2456-2501, 2017.
20. Chai, Y., He, L., Mei, Q., Liu, H. and Xu, L., Deep Learning Through Two-Branch Convolutional Neuron Network for Glaucoma Diagnosis. Proceedings of International Conference on Smart Health, pp. 191-201, Springer, 2017.
21. Zhang, J., Luo, Y., Jiang, Z. and Tang, X., Regression Analysis and Prediction of Mini-Mental State Examination Score in Alzheimer’s Disease Using Multi-granularity Whole-Brain Segmentations. Proceedings of International Conference on Smart Health, pp. 202-213. Springer, 2017.
22. Liu, Y. and Choi, K.S., Using Machine Learning to Diagnose Bacterial Sepsis in the Critically Ill Patients. Proceedings of International Conference on Smart Health, pp. 223-233, Springer, 2017.
23. Viegas, R., Salgado, C.M., Curto, S., Carvalho, J.P., Vieira, S.M. and Finkelstein, S.N., Daily prediction of ICU readmissions using feature engineering and ensemble fuzzy modeling. Expert Systems with Applications, 79, pp.244-253, 2017.
24. M. Saeed, M. Villarroel, A.T. Reisner, G. Clifford, L. Lehman, G.B. Moody, T. Heldt, T.H. Kyaw, B.E. Moody, R.G. Mark. Multiparameter Intelligent Monitoring in Intensive Care II (MIMIC-II): A public-access ICU database. Critical Care Medicine 39(5):952-960; doi: 10.1097/CCM.0b013e31820a92c6, 2011.
25. Dong, Y., Wang, Q., Zhang, Q. and Yang, J., Classification of Cataract Fundus Image Based on Retinal Vascular Information. Proceedings of International Conference on Smart Health, pp. 166-173, Springer, 2016.
26. Zheng, B., Zhang, J., Yoon, S.W., Lam, S.S., Khasawneh, M. and Poranki, S., Predictive modeling of hospital readmissions using metaheuristics and data mining. Expert Systems with Applications, 42(20), pp.7110-7120, 2015.
27. Fialho, A.S., Cismondi, F., Vieira, S.M., Reti, S.R., Sousa, J.M. and Finkelstein, S.N., Data mining using clinical physiology at discharge to predict ICU readmissions. Expert Systems with Applications, 39(18), pp.13158-13165, 2012.