Correlating Health and Wellness Analytics for Personalized Decision Making

Wajahat Ali Khan, Muhammad Idris, Taqdir Ali, Rahman Ali, Shujaat Hussain, Maqbool Hussain, Muhammad Bilal Amin, Asad Masood Khattak, Yuan Weiwei, Muhammad Afzal, Sungyoung Lee, Byeong Ho Kang

Abstract— **Personalized healthcare envisions providing customized treatment and management plans to individuals at their doorstep. Key factors to ensure personalized healthcare is to involve with the individual in their daily life activities and process the gathered information to provide recommendations. We identified the mostly exposed domains for gathering chronic disease patients information that includes: clinical, social media, and daily life activities. Clinical data is related to the health-care of the patients while social media, sensory, and wearables data is related to the wellness data of the patients. A framework is required to monitor the health and wellness information of the patients for health and wellness analytics provisioning to the physicians for better decision making. We propose Personalized, Ubiquitous Life-care Decision Support System (PULSE); a state of the art decision support system that helps physicians and patients in life-style management of chronic disease patients such as Diabetes. The proposed approach not only utilizes clinical information but also personalized information by correlation to find hidden information using big data health analytic for improvement of life-care. PULSE provides health analytics by utilizing and processing clinical information of the patient. In the same way, it provides wellness analytics to the patients by using their social, activities, emotions and daily life information. The co-relation between clinical and personalized analytics is performed for better recommendations to the patients. This eventually results in improved life-care and healthy living of the individuals.**

I. INTRODUCTION

The advent of technology has provided means for physicians and patients to monitor the management of chronic diseases using health analytics. Medical data is continuously growing and daily life activities of a patients has direct effect on their normal and abnormal medical observations. Focus of the healthcare research community is on remotely management of chronic disease patient. This includes support for patient self-management, responsibility shift from clinical providers, and reducing the use of emergency department and hospital services to reduce the overall spending on chronic disease [1]. Management of chronic diseases such as diabetes require health analytics for improvement of life-care. Health analytics is dependent on medical and daily life data collected from Electronic Medial Record (EMR) and sensory technologies.

Clinical and life-care decision support systems mostly focus on either clinical data processing or simply managing the life-care by the use of sensors technologies information. Daily life activities monitored by sensors technologies require clinical information from the EMR as well for better management of the disease. The clinical data can be obtained as patient's EMR that shows encounters information of the patients. The activities between two encounters contains information that can be used for life-care decisions or analysis of which can help the patient improve his/her life style. Monitoring of these activities can be performed by using wearable, accelerometer, GPS, and other types of sensors. A system is required to manage the fusion of initial data acquired from different sensors and then combining both forms of the data for health analysis and proper disease management.

We previously developed systems that achieve clinical and wellness recommendations objectives; with Smart CDSS [2] purely processing clinical data and personalized data using health-care standards for providing clinical guidelines and recommendations to the patients and clinicians, and ATHENA [7] system curates sensory data for wellness services delivery. Smart CDSS was mainly focused on the integration of sensory and clinical data, and processing the integrated information using Medical Logic Modules (MLM) based clinical knowledge bases. We faced number of challenges for the integration of clinical and sensory data. Firstly, the knowledge base creation considering health and wellness information is mainly dependent on the domain experts. Secondly, standard based reasoning of integrated health (reliable) and wellness (less reliable) information. Thirdly, clinical data is more reliable while sensory data reliability is at the lower level, therefore decision making based on both sets of data was somewhat risky.

To overcome these challenges we come across the idea of PULSE, that correlates the clinical and sensory data rather than integrating them. The integration was at the data level and therefore plenty of challenges were faced, while

W.A.Khan, M.Idris, T.Ali, R.Ali, S.Hussain, M.Hussain, M.B.Amin, M.Afzal, and S.Y.Lee are with Department of Computer Engineering, Kyung Hee University, Seocheon-dong, Giheung-gu, Yongin-si, Gyeonggi-do, Republic of Korea, 446-701, {wajahat.alikhan, idris, taqdir.ali, rahmanali, shujaat.hussain, maqbool.hussain, mbilalamin, muhammad.afzal, sylee}@oslab.khu.ac.kr

A.M. Khattak is with College of Technological Innovation, Zayed University, Abu Dhabi, UAE, asad.khattak@zu.ac.ae

W.Yuan is with Department of Computer Science and Technology, Harbin Engineering University, Nantong Street 145, 150001, Harbin City, China, yuanweiwei@khu.ac.kr

B.H.Kang is with School of Computing and Information Systems, University of Tasmania, Australia Byeong.Kang@utas.edu.au

correlating works at the service level. Clinical and Sensory Knowledge Bases works separately to processes the heterogeneous data separately. Clinical data is converted to health-care standard vMR, and then processed using MLM for clinical recommendations services creation; while sensory data is processed using machine learning based reasoning for generating the personalized recommendations services. Correlation among the services generated are performed by the experts.

The proposed solution, PULSE is a life-care decision support system that provides health analytics services in the form of monitoring clinical, personalized and feedback information of the chronic disease patients. Unlike Smart CDSS that converts both forms of the data of the patient into Virtual Medical Record (vMR) standard and then processing the data based on Medical Logic Module (MLM) oriented rules in the knowledge base; PULSE processes clinical knowledge with the same standard based pattern while personalized knowledge with machine learning based reasoning techniques. Big data techniques are used to manage the data coming from different heterogeneous sources and converting them into structured format for processing; and feedback analysis is performed by taking the consent of the patients about the recommendations or analysis provided to them by the proposed system. A case study is presented to describe PULSE system prototype[1] developed for providing health and wellness analytics in Section V.

II. Motivation

Future directions of health-care community focus on progress from learning health systems to person-centered care delivery [18]. The person-centered care delivery envisions shareable and integrated EHR, with prime importance to the role of social media information and also information obtained using Internet of Things (IoTs). Reliability is the factor that differentiates the EHR and IoTs data, therefore, leading to challenges of integrating or correlating both types of data in health-care and wellness applications. Figure 1 provides the spectrum for the requirement of a framework to correlate health and wellness applications. EHR/EHR//PHR are some of the sources to obtain health related data of patients or users. IoTs that includes sensors, wearables, and social media data contributes to the wellness information of the users. These sources generates big data for personalization aspect to effectively utilize this information. Intelligent processing using health-care standard based inferencing and machine learning based reasoning approaches will contribute in health and wellness analytics respectively. As previously mentioned, reliability of health data sources is more as compare to wellness sources, therefore correlating both the information will identify hidden patterns for better decision making.

The two aspects of integration and correlation of health and wellness data is a challenging task and require different system design and implementation aspects. The comparison between both the approaches is described in Table I.

[1]http://uclab.khu.ac.kr/PULSE//

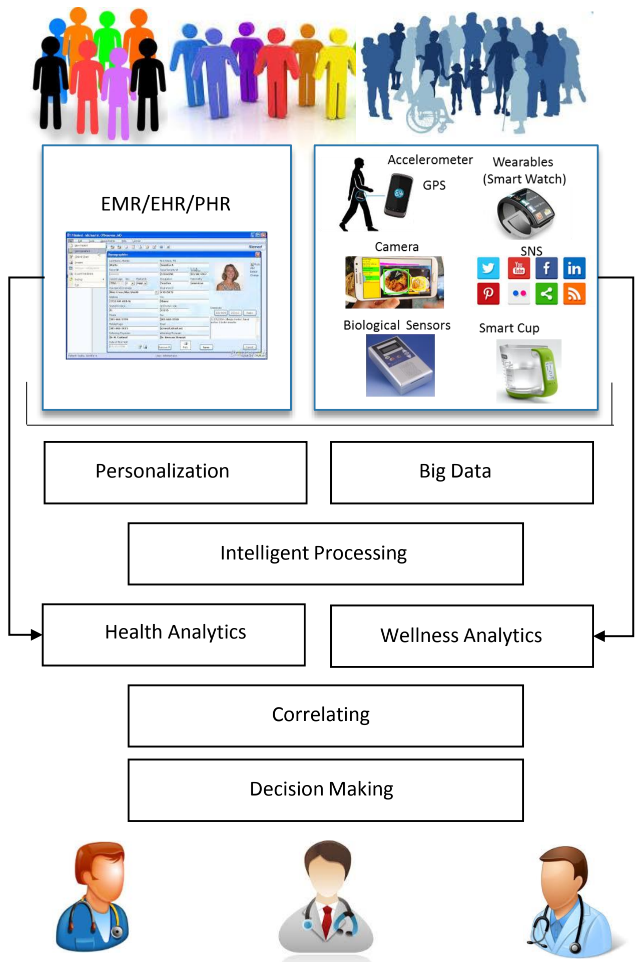

Fig. 1. Internet of Things (IoTs) and Decision Making Spectrum

TABLE I
Comparison of Integration and Correlation of Health and Wellness Information

Integrating	**Correlating**
A combined knowledge base of health and wellness	Separate knowledge bases for health and wellness
Single reasoner for inferencing on integrated data	Multiple reasoners for health and wellness data
Integrating reliable clinical data with less reliable sensory data increases the risks in decision making, as the veracity aspect of big data is maximum in sensory data	Correlating health and wellness information at service level
Decision making on integrated data and knowledge bases, thus single recommendation	Health and Wellness Analytic for correlating recommendations and finding the hidden information

Comparison demonstrates the effectiveness of correlation of health and wellness information compared to integration.

Therefore, a framework to support correlation of health and wellness information is required for better decision making. This paper presents a novel framework that support such claims and is called as PULSE.

III. RELATED WORK

Existing systems are mainly clinical decision support systems that focus on alerts, reminders and recommendations based on only clinical data. Vermont Diabetes Information System [3] is a decision support system based on the chronic care model targeting providers and patients. It is designed for easy integration into primary care offices with or without EMR. It uses e-mail reminders when testing was overdue or results elevated and is only for low disease severity in participants. It's main drawbacks includes its only focus on EMR based data monitoring, reminder based system with no recommendations, and targets only low disease severity patients. Diabetes Wizard [4] is an EHR based Diabetes CDSS with main focus on Glucose and Blood Pressure Control. Nurse and physician fills the diabetes wizard form in patient encounter and finally recommendation is provided for medications change, follow up care and treatment. Its focus is also only on EMR data and no personalized services and monitoring of patients activities and his behavior analysis is provided to the patients. Diabetes Manager Clinical Decision Support System (CDSS) [5] facilitates the management of diabetes by primary care providers. It empowers patients and encourages self-management of their diabetes by interventions on diabetes related clinical information. Patient has to regularly upload their information. The main drawbacks of the system includes burden on patients to upload data, not regular diabetes patient monitoring, and focus on only unreliable patient data entry. CDS Starter Kit [6] is another decision support system in which initially patient reason for visit is recorded. Afterwards measuring and recording vitals, current medications, and other pertinent data are performed. Finally follow up care and reviewing tests, vital signs and foot exam is conducted. The main drawbacks of this system are it's a reminder system for follow up care only, focus only on EMR data, no personalized services and no regular monitoring of patients.

IV. PULSE FRAMEWORK

The abstract view of layered architecture of PULSE framework is shown in Figure 2 and its component details is presented in the subsequent sections. The components are developed and their corresponding papers are referred, in detail of each component, for insight into the comprehensive view.

- ***Data Acquisition Layer*** layer consists of multimodal sensors data such as accelerometer, magnetometer, gyroscope, proximity sensor, ambient light sensor, GPS and cameras in smartphones and wearable sensors. The social networks behaves as logical sensors and its data as well as clinical data from the EMR is also passed to big data storage and processing layer.
- ***Big Data Storage and Processing Layer*** process the clinical and sensory data obtained from the *Data Acquisition Layer*. The unstructured sensory data and semi structured clinical data is stored in the Hadoop distributed file system (HDFS) [9]. The objective of using hadoop for the proposed platform is to provide scalability and fault tolerance to the bulk amount of everyday generated sensory logs and clinical data. The retrieved data from HDFS is further stored into the *Personalized Intermediate Data* for further processing and providing fast access to the subcomponents [8]. We use Apache Pig scripting language[2] and ZooKeeper[3] technologies to fulfill the data intensive approach.
- ***Personalized and Clinical Data Processing Layer*** is responsible for the intermediate processing of the variety of data by the corresponding components. *Social Media Interactor* module aim is to improve the patients health and lifestyle by utilizing his/her social interaction based on different social networks. For instance, after observing a patients daily social media activities, SMIE (Social Media Interaction Engine) finds some complications with his/her lifestyle like he/she usually sleeps late; does not exercise regularly; does not take medicine on time; eats too much. Obviously, these lifestyles are not good for chronic disease patients and are identified by this module [10]. *Activity and Emotion Recognizer* consists of the multimodal sensor based activity and emotion recognition modules. The outputs of these individual independent modules are fused together to predict the real time high level contexts. It consists of both video and audio emotion recognition as well as activity recognition [11] [12]. *Human Behavior Analyzer* helps in understanding the human behavior which can prove helpful in supporting an active lifestyle and wellbeing [8]. This module takes the low-level and high-level contexts as input and analyzes a subjects short-term and long-term behavior. *Context Aware Recognizer* performs reasoning on high level activities from low level activities recognized by different sensors. It uses ontology for modeling the context, user profile information, and representation of information [13]. *Health Information Manager*takes the clinical information from EMR in the form of clinical records. It extracts important clinical facts about the patient/user to from the record for performing reasoning to provide clinical recommendations. It provides this information to the *Interoperability Adapter [14].*
- ***Adapters Layer*** provides the interface through which data is collected and provided to the corresponding modules. Also, the processed information is forwarded to the reasoning modules through these adapters. The *Interoperability Adapter* is responsible for converting the extracted clinical information into vMR standard

[2]https://pig.apache.org/ (Last Visited March 5, 2015)

[3]http://www-01.ibm.com/software/data/infosphere/hadoop/zookeeper/ (Last Visited March 5, 2015)

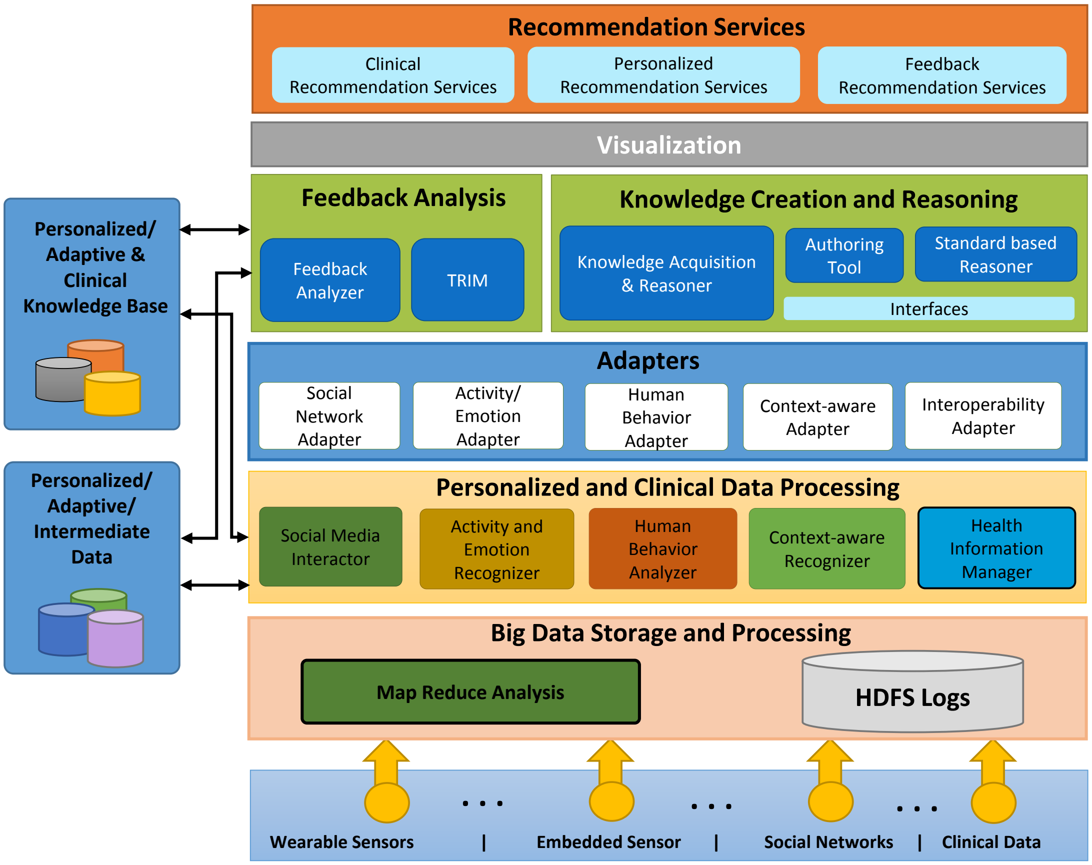

Fig. 2. PULSE Framework

format.

- ***Knowledge Creation and Reasoning Layer*** coordinates the overall reasoning process on sensory and clinical data. The *Knowledge Acquisition and Reasoner* [16] access all the sensory and social media related information from the adapters except *Interoperability Adapter*. It aims to provide personalized and automatically acquire knowledge from vast amount of data by using different machine learning approaches. *Authoring Tool* [15] provides the facility to physicians to enter their knowledge into the knowledge base that is transformed to rules for recommendations to be generated. The knowledge is stored in the form of Medical Logic Module (MLM). The *Standard based Reasoner* performs vMR based reasoning on clinical datasets. It also supports evolution and maintenance of clinical knowledge bases.

  Feedback Analysis is composed of two kinds of feedback. The *Feedback Analyzer* module deals with the user satisfaction based on the recommendation provided to the user. On the other hand, TRIM monitors the feedback of other users to provides this information to the respective user. If the user finds another users feedback on recommendations is continuously helpful, he/she can give explicit feedback on the trustworthiness of this user. This trustworthiness is computed by the user trust instead of user similarity and the search within the trusted network of PULSE subscribed users is performed by a novel technique S_Search [17], which search the whole network and only recommends the most trusted users.

  The *Knowledge Correlation Module* is responsible for correlated the clinical and personalized recommendation services and utilize *Visualization* component for providing the services to the users. *Visualization* component shows users clinical and personal information in different graph templates to give him/her a different insights like his activities information and his diabetes level in the past week.

V. CASE STUDY

Pulse maintains and monitors the medical record and physical activities of a patient and provides graphical analysis to understand the pattern and improve the patients health. Considering an example of a patient that subscribed to PULSE system for monitoring his health and activities. The different encounters data related to diabetes (such as Diastolic Blood Pressure(DBP), Systolic Blood Pressure (SBP), Fasting Blood Sugar (FBS), Total Cholesterol (TC) and others) is shown in Figure 3. His physical activities during the encounters are monitored using the wearable sensory data. Based on this dataset, the relation among exercise,

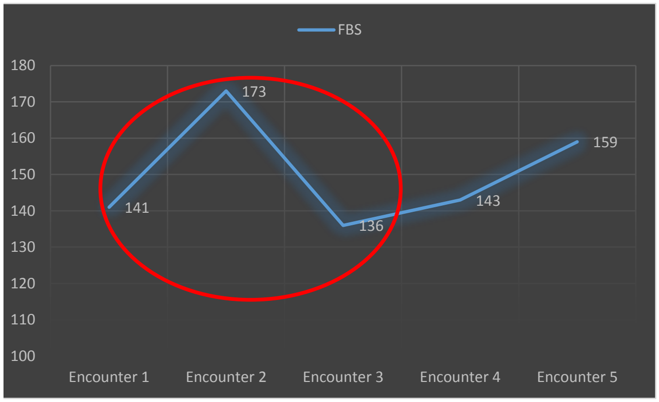

Fig. 4. FBS value in different Encounters

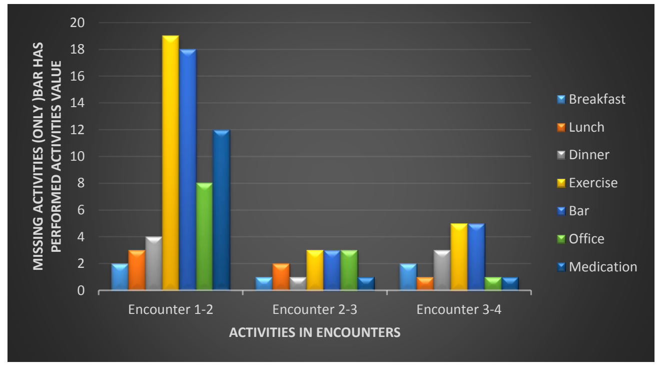

Fig. 5. Activities Information in different Encounters

diet, medical readings, and action suggestions are graphically interpreted and shown to the user/patient for further action.

Medical history, as shown in Figure 3, represents the clinical readings for Blood Pressure (Diastolic, Systolic), PP2H, TC, TG, LDL, and FBS in particular. Based on the normal (healthy person) statistics, each attribute is colored from Red (action suggested) to Green (Excellently maintained). These statistics individually represent patients records and gives a sign of action or satisfaction. However, to explicitly understand the reason and duration behind any specific medical readings such as FBS, these readings are represented graphically as in Figure 4. Here, between two consecutive encounters, the patient or doctor can easily understand the pattern and downfall or rise in any of the measured stats. In Figure 4, FBS between encounter 1 and 2 is highly increased to abnormal state based on the data in Figure 3. Similarly, between other encounters, it has been controlled. These types of graphs are presented for each statistics and the user can select any combination of the medical readings such as FBS between encounters from 1 to 5 or FBS, DBP, SBP between encounter 2 and 3, and others.

Health analytics can help in identifying the abnormal observations and finding out the personalized analytics in specific encounters duration can help in improving the lifestyle and better management of the chronic disease. The daily life activities monitored are also depicted by different colors that shows abnormalities in the different activities in Figure 5. The activities monitored includes: Breakfast, Lunch, Dinner, Exercise, Bar, Office and Medication. Bar activities includes the number of time patient has visited the bar while other activities shows the number of times these are missed by the patient. Visualization of these shows that medication, exercise and bar are the activities that are not performed properly by the patient and therefore, encounter 2 shows high FBS value.

VI. CONCLUSION

The proposed system provides the infrastructure to support the health and wellness analytics, by investigating the big data collected from heterogeneous sources, to facilitate physicians or experts in hidden information identification for better decision making. The objective is achieved by correlating the personalized sensory and clinical data with analytic schemes. The services are provided by constructing intelligent knowledge base using social, sensory and clinical information.

ACKNOWLEDGMENT

This work was supported by the Industrial Core Technology Development Program (10049079 , Develop of mining core technology exploiting personal big data) funded by the Ministry of Trade, Industry and Energy (MOTIE, Korea)" and was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIP) NRF-2014R1A2A2A01003914

REFERENCES

[1] Coye, M. J., Haselkorn, A., & DeMello, S. (2009). Remote patient management: technology-enabled innovation and evolving business models for chronic disease care. Health Affairs, 28(1), 126-135.
[2] Hussain, M., Khattak, A. M., Khan, W. A., Fatima, I., Amin, M. B., Pervez, Z., ... & Latif, K. (2013). Cloud-based Smart CDSS for chronic diseases. Health and Technology, 3(2), 153-175.
[3] MacLean, C. D., Gagnon, M., Callas, P., & Littenberg, B. (2009). The Vermont diabetes information system: a cluster randomized trial of a population based decision support system. Journal of general internal medicine, 24(12), 1303-1310.
[4] OConnor, P. J., Sperl-Hillen, J. M., Rush, W. A., Johnson, P. E., Amundson, G. H., Asche, S. E., ... & Gilmer, T. P. (2011). Impact of electronic health record clinical decision support on diabetes care: a randomized trial. The Annals of Family Medicine, 9(1), 12-21.
[5] Yolande Calhoun, Paul Giam, Lindsay Kaplan, Biljana Spasic, "Diabetes Manager CDSS: A Diabetes Management Tool for PCPs and Patients", *Technical Report*. http://162389445480246740.weebly.com/uploads/1/4/5/9/14590780/final_project.pdf [Last Visited May 13, 2013]
[6] CDS Starter Kit: Diabetes follow-up care, http://www.healthit.gov/sites/default/files/del-3-7-condition-example-diabetes.pdf [Last Visited May 13, 2013]
[7] Fahim, M., Idris, M., Ali, R., Nugent, C., Kang, B., Huh, E. N., & Lee, S. (2014). ATHENA: A Personalized Platform to Promote an Active Lifestyle and Wellbeing Based on Physical, Mental and Social Health Primitives. Sensors, 14(5), 9313-9329.
[8] Hussain, S., Kang, B. H., & Lee, S. (2014). A Wearable Device-Based Personalized Big Data Analysis Model. In Ubiquitous Computing and Ambient Intelligence. Personalisation and User Adapted Services (pp. 236-242). Springer International Publishing.
[9] Shvachko, K., Kuang, H., Radia, S., & Chansler, R. (2010, May). The hadoop distributed file system. In Mass Storage Systems and Technologies (MSST), 2010 IEEE 26th Symposium on (pp. 1-10). IEEE.

	EncounterDate	Weight (kg)	Height (cm)	SBP	DBP	FBS	PP2h (mg/dl)	HBA1c (%)	TC	TG	LDL	HDL	Cr (mg/dl)	Hypoglycaemia
Encounter 5	5/7/2011	59	170	112	64	159	Null	6.7	206	42	90	85	1.05	No
Encounter 4	5/6/2011	58	170	120	70	143	293	5.9	201	47	99	90	1.01	No
Encounter 3	5/5/2011	59	170	117	56	136	Null	6.9	209	24	108	85	1.04	No
Encounter 2	5/4/2011	60	175	114	68	173	Null	7	206	64	112	75	0.92	No
Encounter 1	5/3/2011	60	175	104	62	141	Null	6.3	192	34	94	74	1.12	No
				Legends										
		Excellent		Good		Action Suggested		Borderline High						

Fig. 3. Diabetes Data in different Encounters

[10] Fatima, I., Fahim, M., Guan, D., Lee, Y. K., & Lee, S. (2011, February). Socially interactive CDSS for u-life care. In Proceedings of the 5th International Conference on Ubiquitous Information Management and Communication (p. 95). ACM.

[11] Siddiqi, M. H., Ali, R., Khan, A. M., Park, Y. T., & Lee, S. (2015). Human Facial Expression Recognition Using Stepwise Linear Discriminant Analysis and Hidden Conditional Random Fields. Image Processing, IEEE Transactions on, 24(4), 1386-1398.

[12] Fahim, M.; Fatima, I.; Lee, S.; Lee, Y.-K. Eem: Evolutionary ensembles model for activity recognition in smart homes. Applied Intelligence 2012, 38, 88-98.

[13] Khattak, A. M., Hung, D. V., Truc, P. T. H., Hung, L. X., Guan, D., Pervez, Z., ... & Lee, Y. K. (2010, July). Context-aware human activity recognition and decision making. In e-Health Networking Applications and Services (Healthcom), 2010 12th IEEE International Conference on (pp. 112-118). IEEE.

[14] Khan, W. A., Khattak, A. M., Hussain, M., Amin, M. B., Afzal, M., Nugent, C., & Lee, S. (2014). An Adaptive Semantic based Mediation System for Data Interoperability among Health Information Systems. Journal of medical systems, 38(8), 1-18.

[15] Ali, T., Hussain, M., Ali Khan, W., Afzal, M., & Lee, S. (2013, July). Authoring tool: Acquiring sharable knowledge for smart CDSS. In Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE (pp. 1278-1281). IEEE.

[16] Ali, R., Siddiqi, M. H., Lee, S., & Kang, B. H. (2015). KARE: A Hybrid Reasoning Approach for Promoting Active Lifestyle.

[17] Yuan, W., Guan, D., Shu, L., & Niu, J. (2014). Recommender Searching Mechanism for Trust-Aware Recommender Systems in Internet of Things. AUTOMATIKA: casopis za automatiku, mjerenje, elektroniku, racunarstvo i komunikacije, 54(4), 427-437.

[18] Payne, T. H., Corley, S., Cullen, T. A., Gandhi, T. K., Harrington, L., Kuperman, G. J., ... & Zaroukian, M. H. (2015). Report of the AMIA EHR 2020 Task Force on the Status and Future Direction of EHRs. Journal of the American Medical Informatics Association, ocv066.