Diabetic Patients Monitoring and Data Classification Using IoT Application

Amine Rghioui[1], Assia Naja[2,3], Abedlmajid Oumnad[1]
[1]Research Team in Smart Communications-ERSC (formerly known as LEC) – Research Centre E3S, EMI.
[2]Mohamed V University in Rabat
[3]International University of Rabat, FIL, TICLab, Morocco
rghioui.amine@gmail.com, naja.assia@gmail.com, aoumnad@emi.ac.ma

***Abstract*—Internet of Things (IoT) is used to implement several applications especially with the increasing number of available data. IoT can be applied in many fields such as patient health monitoring system. For instance, we can use IoT for the field of health in a way to analyze the data and present it to doctors and paramedical staff. By understanding, processing and utilizing the knowledge and information hidden in Big Data with respect to health issues and disease trends in a certain population, we can find solutions, to live longer and healthier. Big data analytics improves health care insights in many aspects. Diabetes is one of the most commonly found health problem in the world. We believe that an early detection of Diabetes aids in better diagnosis and treatment For that we intend to realize a system that can automatically transmit the blood glucose measurements via Short Message Service (SMS) addressed to the doctor. The results then showed the validity of the test in the form of validation of sensor measurement data and output data SMS.**

***Index Terms*—Internet of Things, Machine Learning, Healthcare, Diabetes, Blood Glucose.**

I. Introduction

Technologies and application of the Internet of Things are having a tremendous advance in important IoT technical fields, becoming each day more accessible, available and versatile, allowing faster growth of the things interconnected through the Internet. The quality of life of patients is an important target of smart Healthcare, and is widespread concern with the construction of new application in this field. The daily mobile healthcare service becomes more and more important. Chronic diseases influence the health of the people living, such as cardiovascular and diabetic diseases.

Patients are also more open to giving away part of their privacy if this could save their lives or other people's lives. "If having more information about me enables me to live longer and be healthier", said Marc Rotenberg of the Electronic Privacy Information Center," then I don't think I'm going to object. If having more information about me means my insurance rates go up, or that I'm excluded from coverage, or other things that might adversely affect me, then I may have a good reason to be concerned" [1].

Diabetic is a dangerous disease that can affect humans, which requires much attention to keep them healthy. This disease becomes one of the main causes of death in the world, and most people are still not aware of the importance of maintaining diabetic health. Diabetes is caused by the increase or decrease in the level of glucose in the blood, it is necessary to keep it stable in a precise interval – a condition where; the level of glucose exceeds a normal level. Some patients with anomaly require intensive care and monitoring to avoid an aggravation [2].

The growth of diabetic patients increased the use of continuous glucose monitoring devices (CGM) which are becoming the new method of continuous monitoring. They provide real-time information about the glucose level which is updated every five minutes. As a result, patients and physicians must treat approximately 9000 readings [3] and interpret the enormous amount of data to adjust insulin doses and keep blood glucose as close to normal as possible [4].

A glucose level concentration value $\leq$70 mg/dl is a hypoglycemic and a glucose concentration value $\geq$180 mg/dl is hyperglycemic. Therefore, it is necessary to make a periodic check for diabetic patients. With the reason of busyness, such as checks, and consultations, however, are sometimes difficult to be done by the patient. In the purpose of consultation, a mobile blood glycemic device is required to send the glucose value to the medical experts. In other words, it takes a Tele-monitoring application to resolve the issue. One of the communication media for tele monitoring applications is SMS services via the GSM network that can be widely accessed.

We will develop a smartphone-based wireless blood glucose monitoring system-using Bluetooth. The measurements of blood glucose will be sent to the mobile phone via Bluetooth and will be shared on the Android application. This paper aims to discuss the realization of digital Blood glucose meter system integrated with GSM network for monitoring a patient's glycemic in real time. The patient can do the measurements by himself/herself and the device will automatically send the data to medical experts via SMS. In our research, we developed a blood glucose measurement system based on E-Health module. E-Health module was integrated in the Arduino to read the measurement data and then to be sent via short message service using the SIM module. This environment is expected to help patient, particularly outpatients, to be able to establish a communication with the medical expert periodically and for some emergency cases, real-time applications are required. Thus, the medical expert still could monitor patients' health

978-1-7281-4341-5/20/$31.00 ©2020 IEEE

remotely and provide early treatment if something happens to the patient.

Health-care data analytics system can help one of the issues that is to concentrate on the cost of patients for individual treatment. Various big data skills load up and investigate more than health care collective by means of good organization cost investments are explained in better healthcare [5]. Many researchers have developed and implemented various analysis and prediction models using different data mining techniques. The authors in [6] proposed the Hadoop and MapReduce based approach for the analysis of diabetic data. In [7], the authors use a classification technique to find out patterns from the diabetes data sets. They employed naive Bayes and Decision Tree algorithms by using Weka tool. For the analysis of diabetic data, many authors preferred the decision tree for the classification, rules generation, pattern recognition etc. [8]. In this work, we are using Naïve bayes, Random Tree, ZeroR, and J48 classification algorithm from the diabetes data set, to test the most powerful to determine the patient's level of risk.

The main contents of this paper are organized follows; Section II describes the patient monitoring elements using in this paper. Section III presents a description of the implementation system. Section IV presents data classification for diabetic patients. Section V provides a brief description of the result discussed. Finally, a conclusion is given in Section VI.

II. PATIENT MONITORING ELEMENTS

A. *Digital Blood Glucose Meter.*

Testing blood sugar with a blood glucose meter is one of the key skills of successful diabetes management. Each individual requires a different blood glucose testing approach for their diabetes, and some will need to test blood glucose more regularly than others. The measurement of blood Glucose is a measure of the level of glycemic in the blood using an electronic device called a blood glucose meter. To do this, you prick your fingertip with a small needle and place a drop of blood on a test strip. The strip is inserted into the blood glucose meter. The digital display shows your blood sugar level shortly afterward.

There is a wide variety of blood glucose meters on the market. Some meters are designed for ease of use; others for ease of transport or connectivity, still more incorporate advanced technology such as USB software or multi-strip testing. [9]

Fig.1. Digital Blood Glucose

B. *E-Health Sensor Shield*

E-Health sensor shield is a special module designed to meet the needs of medical device design based on Arduino. It provides information collected from 9 different biometric sensors. The biometric e-Health Sensor Shield v2.0 can therefore, be used for either monitoring in real-time or gathering data for later analysis. The data collected can be transmitted via the shield's various interfaces (Wi-Fi, GPRS, Bluetooth, and 3G, 802.15.4 or ZigBee). It is used for a communication interface between the sensors (BP meter) with Arduino. The figure2 shows the E-Health sensor.

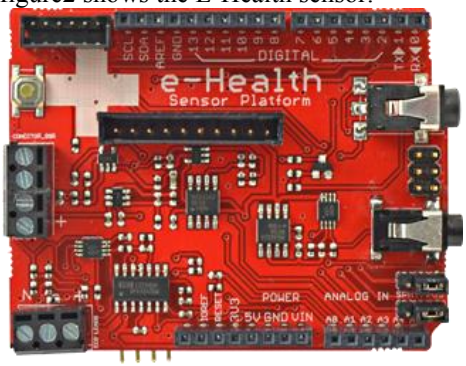

Fig.2. E-Health Shield [6]

C. *GSM Modem*

In this research, we used Wavecom FastTrack M1306B series modem as an SMS sender module. Wavecom modem enables Arduino board connected to a mobile network to send and receive SMS either in the form of text or in the PDU format. Basically, the process of sending and receiving SMS on Wavecom modem used AT command.

Fig.3. GSM Modem

D. *Arduino Card*

Arduino is an open-source computer hardware and software company, the project, and user community that designs and manufactures single-board microcontrollers and microcontroller kits for building digital devices and interactive objects that can sense and control objects in the physical world. Arduino board designs use a variety of microprocessors and controllers. The boards are equipped with sets of digital and analog input/output (I/O) pins that may be interfaced to various expansion boards (shields) and other circuits. The boards feature serial communications interfaces, including Universal Serial Bus (USB) on some models, which are also used for loading programs from personal computers.

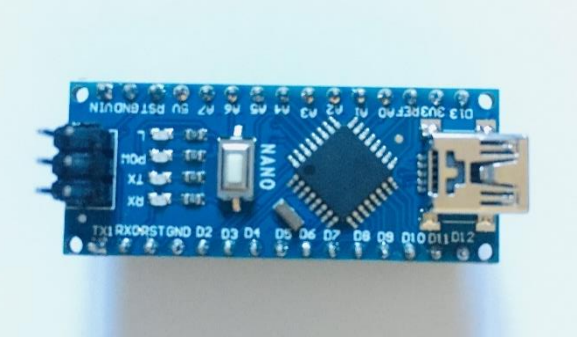

Fig.4. Arduino Card

III. IMPLEMENTATION

This section describes the implementation of the system in details including the scheme, system illustration, hardware, and software installation.

The flowchart of the system in Figure 5 shows the complete system from the measurement of glucose level in the patient until the final decision of the doctor.

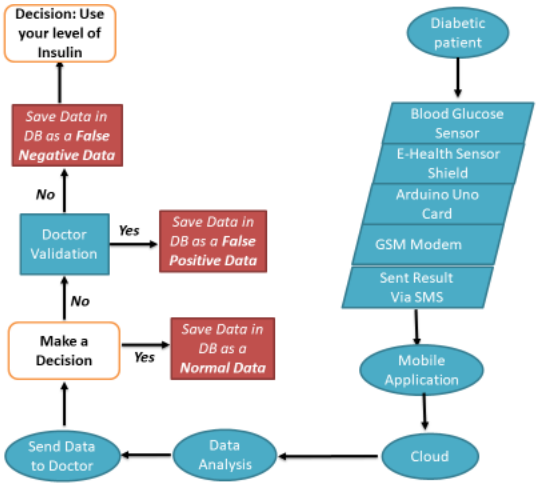

Fig.5. Flow chart of the system

In this system, the patient can perform a blood sugar level measurement using the Digital Blood Glucose Meter. After the measurement, he will get the results of blood sugar level. The measurement data is automatically transmitted by the sensor in serial mode to Arduino and it will be processed by the microcontroller. Data processed by a microcontroller would be automatically transmitted to the GSM modem that has been integrated into the Arduino board to send the measurement results to a specialist doctor via SMS.

A. Hardware Installation

The overall design of the system was accomplished by connecting the Blood Glucose Level Sensor to E-Health Shield, which was connected to the Arduino board. Arduino was also connected to the GSM module with TTL to RS232 converter. Data rates for serial communication between the microcontroller, the sensor and the GSM module was 115200 bps. Data is formatted to create synchronization between the sender and the recipient. The hardware installation can be seen in Fig. 6

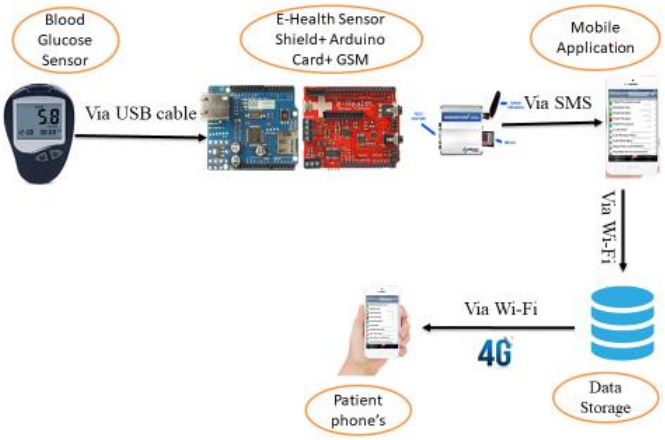

Fig.6. Hardware Implementation

Testing the overall system was accomplished by integrating the entire circuit or the configuration of all modules required in the design of this system. Blood glucose sensor needed to be connected to the E-Health Shield, which has been connected to the Arduino board. At the same time, Arduino should also be connected to the GSM module. For the serial communication between the microcontroller, sensor and GSM module, they were set at the same rate. In addition, the type of data being transmitted in the program must also be set to synchronize transmitter and receiver.

B. Software Installation

The measurement data from the sensors would be sent via the serial data to Arduino Shield E-health. The measurement data was displayed via the SMS service in accordance with the authentication recipient. The listing program of data transmission using the SMS based on AT-Command is presented as follows:

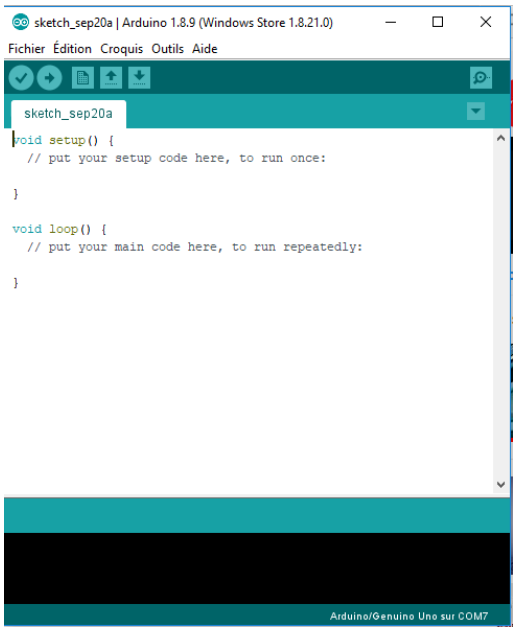

Fig.7. Arduino Software interface

IV. Classification Data in Healthcare

Classification is the process of converting the data records into a set of classes. Classification involves predicting a certain outcome based on a given input. In order to predict the results, it needs to fetch the data already available.

There are several classification models; namely, decision tree, ripper rule, neural networks, naive bayaïs, k-nearest neighbors and support vector machine. In this section, we will present the decision tree algorithms, naive bayaïs algorithms, Random Forest algorithms, and neural networks.

1) Decision Tree

A decision tree is a widely used classification technique. There are several related works in the literature about classification data with the decision Tree algorithm, Milan Kumari.et al [10] used decision tree to predict cardiovascular disease and they classify the patients who have swollen glands and diagnose the patients whether they are infected from fever, cold or throat pain. AI-Radaideh, et al [11] applied a decision tree method to predict the grade of the student. Even though various classification techniques were applied on the same, decision tree gives the better result.

2) Naïve Bayes

A naive Bayes classifier is a probabilistic classifier belonging to the family of linear classifiers (whose role is to classify in groups. The approach used in Naïve Bayes classifier is very simple. With the help of a small amount of training data, it is possible to classify the given instances [12].

3) Random Forest

Random Forest is a supervised learning algorithm, it creates a forest and makes it somehow random. It is a non-parametric but powerful statistical method which allows to take into account regression problems as well as classification problems with two and several classes, in a single and versatile framework. One of the great advantages of the random forest is that it can be used for both classification and regression problems, which make up the majority of today's machine learning systems. Random Forest has almost the same hyper parameters as a decision tree or bagging classifier

4) Neural Networks

A neural network is expressed in terms of biological neuron system. It consists of number of separate units. It is similar to the brain composed of many processing components, it has a distinctive feature as a three-layer feed-forward neural network: an input layer, a hidden layer, and an output layer. S.Gopika [13] used the neural networks algorithm for diagnosing the renal disease.

5) ZeroR

ZeroR is the simplest classification method that relies on the target and ignores all predictors. ZeroR classifier simply predicts the majority category (class). Although there is no predictability power in ZeroR, it is useful for determining a baseline performance as a benchmark for other classification methods.

6) J48

is an algorithm used to generate a decision tree. The decision trees generated by J48 can be used for classification, and for this reason, J48 is often referred to as a statistical classifier. The objective of this type of method is to build a classification function that can be represented by a tree that is built starting from the root and going towards the leaves.

V. Result and Discussion

In this section, we discuss the results of the testing system. The analysis is focused on the blood Glucose data communication device as well as the performance of communication with Arduino GSM modem to send a short message service.

The plan for this research is to evaluate the performance of methods for classifying diabetes data. The data is evaluated using the J48, Naïve Bayes, RandolTree and ZeroR algorithms. We used a dataset that included 10 patients, for a total of 40 days, 7 men and 3 women aged 27 to 60 years. This research work deals mainly with the accuracy of classification algorithms with respect to execution time and error rate with WEKA software[14]. The results of various measurements are given in Table 2.

From the test results as shown in Table 1, it can be stated that the measurement data was successfully sent. When referring to the GSM communication the test results only showing text messages can show a waiting time difference or delay between sending and receiving SMS. Among 24 times of testing, the range of delay between transmitting and receiving SMS was

TABLE I
Glucose Measurement

Day	Morning	Evening
1	1.34	0.78
2	1.00	0.78
3	0.99	1.34
4	1.32	1.03
5	1.09	1.01
6	1.34	1.22
7	0.89	0.87
8	1.22	1.15
9	1.41	1.33
10	1.06	0.94

from 33 to 122 seconds, and the average was 48.27 seconds. This result was due to the relatively slow data processing in the microcontroller as it must acquire data from sensors in advance before being transformed into a short message format and sent to the medical side. Figure 8 shows a graph of the measurement of the level of Glucose in the blood.

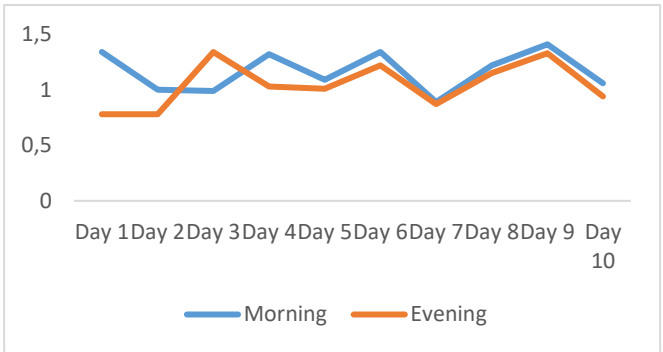

Fig.8. measurement of the level of Glucose in the blood

The dataset during this work is tested and analyzed with four Classification algorithms that are Naïve Bayes, J48, ZeroR

and RandomTree (using training set). In addition, a comparison of accuracy of all classifiers is done and finally it has been investigated that RandomTree technique performs better with an accuracy of 97.87 %. The accuracy level of all the algorithms is given below in Table 2.

TABLE II
THE ACCURACY LEVEL

Algorithms	Correctly Classified Instances	Incorrectly Classified Instances
Naïve Bayes	91.6301%	8.3699%
RandomTree	97.8712%	2.1298%
ZeroR	59.6078%	40.3922%
J48	94.3464%	5.6536%

From table 2, it can be concluded that the accuracy of classifications algorithms like ZeroR, RandomTree, J48, and NaiveBayes for the correctly classified instances is higher as compared to the incorrectly classified instances. Figure 10 shows the bar graph of the correctly and incorrectly classified instances of the taken algorithms.

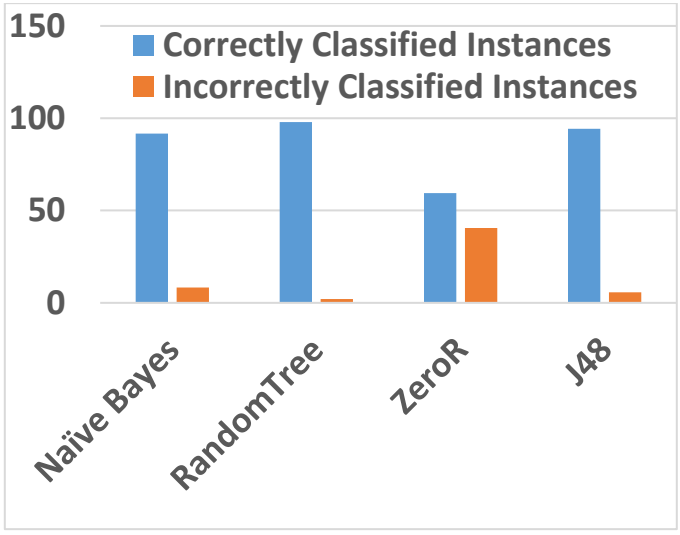

Fig.9. the graph of correctly and incorrectly classified instances of algorithms

Figure 10 shows the time graph of various classification algorithms. The longest time is taken by ZeroR consuming a time of 0.08 seconds and the shortest time is taken by RandomTree consuming 0.03 seconds only.

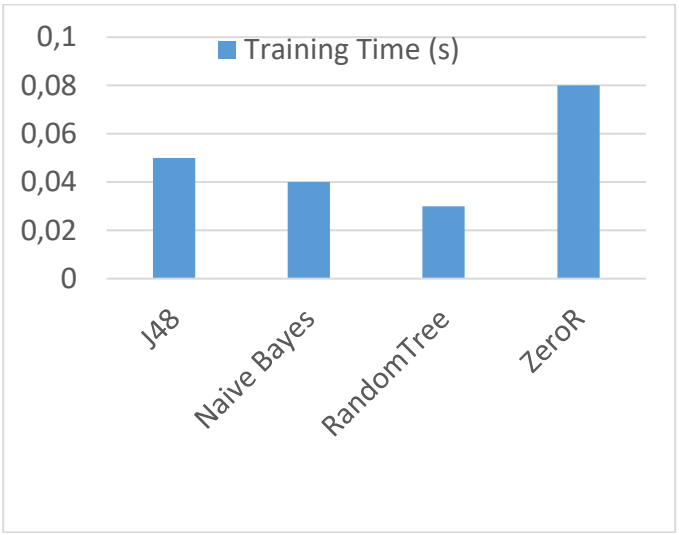

Fig.10. the graph for training time results for top four classifiers

VI. CONCLUSION

The real-time monitoring system for diabetes condition is able to perform with the function of the Internet of Things (IoT). The effectiveness of personal diabetes monitoring via mobile application database has been demonstrated. In this paper, we presented an implementation of the blood Glucose measurement system with GSM module for telemonitoring. The evaluation result showed that the system could work appropriately. The data resulted from measurements were sent and received with a valid value via SMS. The average delay in the transmission was 48.27 seconds. The connection between the blood Glucose sensor and Arduino could not run simultaneously when the data acquisition process by the microcontroller was running. Thus, when a data acquisition process would be performed, the cable between the sensor and the e-health shield must be manually connected. The proposed prediction model has been evaluated by simulation experiments and the results demonstrate that the proposed model improves the prediction accuracy. This model will allow preventive intervention through self-disease management and improve glucose control.

REFERENCES

[1] S. Din, H. Ghayvat, A. Paul, A. Ahmad, M. M. Rathore, I. Shafi, "An architecture to analyze big data in the Internet of Things", 9th International Conference on Sensing Technology (ICST), 2015.

[2] S. A. Tabish, "Is Diabetes Becoming the Biggest Epidemic of the Twenty-first Century ? ", International Journal of Health science, 2007.

[3] E. Aboufadel, R. Castellano, and D. Olson, "Quantification of the variability of continuous glucose monitoring data," *Algorithms*, vol. 4, no. 1, pp. 16–27, 2011.

[4] R. Bunescu, N. Struble, C. Marling, J. Shubrook, and F. Schwartz, "Blood Glucose Level Prediction Using Physiological Models and

Support Vector Regression," *2013 12th Int. Conf. Mach. Learn. Appl.*, vol. 1, pp. 135–140, 2013.

[5] Katon, W. J., Rutter, C., Simon, G., Lin, E. H., Ludman, E., Ciechanowski, P., & Von Korff, M. (2005). The association of comorbid depression with mortality in patients with type 2 diabetes.Diabetes care, 28(11), 2668-2672.

[6] Saravanakumar, Eswari, Sampath, Lavanya "Predictive Methodology for Diabetic Data Analysis in Big Data," ELSEVIER, ISBCC 2015.

[7] Aiswarya Iyer, S. Jeyalatha, Ronak Sumbaly "Diagnosis of Diabetes Using Classification Mining Techniques," IJDKP Vol.5, No.1, January 2015.

[8] Dr.Rajni Jain, "Rule Generation Using Decision Trees," IASRI

[9] https://www.diabetes.co.uk/diabetes_care/blood_glucose_monitor_guide.html.

[10] Cooking-hacks.com, "E-Health Sensor." [Online]. Available: https://www.cooking-hacks.com/documentation/tutorials/ehealthbiometric-sensor-platform-arduino-raspberry-pi-medical#get_kits.

[11] Milan Kumari, Sunila Godara," Comparative Study of Data Mining Classification Methods in Cardiovascular Disease Prediction", IJCST Vol. 2, Iss ue 2, June 2011.

[12] AI-Radaideh, E. W. AI-Shawakfa, and M. I. AI-Najjar, "Mining student data using decision trees", International Arab Conference on Information Technology (ACIT'2006), Yarmouk University, Jordan, 2006.

[13] R. Savundhariyalachmi, N. Pandimeena, P. Ramya," Study of Classification algorithm in Data mining", International Journal of Science and Research

[14] Hall, M., E. Frank, G. Holmes, B. Pfahringer, & P. Reutem. (2009). The WEKA Data Mining Software: An Update. SIGKDD Explorations, 10-18.