Selecting Cloud Service for Healthcare Applications: From Hardware to Cloud Across Machine Learning

Ivan Kuzlo[1], Anastasiia Strielkina[2] [0000-0002-7760-7367], Artem Tetskyi[2] and Dmytro Uzun[2] [0000-0001-5574-550X]

[1, 2] National Aerospace University "KhAI", Kharkiv, Ukraine
[1] i.kuzlo@student.csn.khai.edu
[2] {a.strielkina, a.tetskiy, d.uzun}@csn.khai.edu

Abstract The paper describes the process of creating the Internet of Things (IoT) healthcare applications and selecting an environment to deploy it. Based on research of healthcare application architecture was proposed selecting of cloud service. It draws our attention to complex architecture with using different sources of medical data like external medical databases, medical equipment and wearable medical and non-medical devices. Much attention is given to using machine learning in the process in the detection of health problems. In addition, the paper describes two levels of machine learning: one for detecting single problems with heals and second for predictions complex reports and providing treatment plan based on data from the first level. The main emphasis in choosing of cloud service is made on scalability and the ability to create multiple neural networks for processing data.

Keywords: Internet of Things, Cloud, Healthcare, Machine Learning.

1 Introduction

1.1 Motivation

Machine learning has come a long way from its early roots in classical mathematic and statistics. Today's machine learning uses analytic models and algorithms that learn from data, finding hidden insights without being explicitly programmed where to look. Using algorithms that learn by looking at thousands or millions of data samples, computers can make predictions based on these learned experiences to solve the same problem in new situations. And they are doing it with a level of accuracy that is beginning to mimic human intelligence [1].

Due to the statistical data[2], the number of devices connected to the Internet extremely grows. Industrial giants report that by 2020 we will have from 21 billion [3] to 30 billion [4] devices connected to the Internet (e.g., car, bicycle, smartphone, watch, fitness tracker or some personal medical devices like tonometer or thermometer). All these smart devices form an Internet of Things (IoT). Many of these devices already targeted to medical usage [5], some of them on first look targeted to other purposes but also can get some medical parameters. In future to make full outpatient study it will be enough to wear special fitness tracker for a day, and all information from it will be automatically sent to doctor or healthcare cloud application where neural network will perform as a doctor.

Machine learning (ML) provides many advantages for healthcare application and IoT. It can help to find hidden pathologies during the medical survey and to add an additional level to avoid mistakes of a doctor. And the biggest advantage is that such systems are able to join medical data from different sources like medical database from few places (like different hospitals), home medical kit (tonometers) and IoT devices (smartwatch, glucose monitors, Holter monitor [6]).

1.2 State-of-the-Art

The idea of creating healthcare systems with machine learning is not new. Such ideas can be separated into two big categories: healthcare research support systems and healthcare decision support systems.

The first type targeted to supporting different types of research centres and help them to store and process big amounts of data. Customers of such solutions types typically make researches in genomic/cancer where they have petabytes of data from research and they need a system to process them. One of the examples of such type of system is Google Healthcare & Life Sciences Solutions [7]. It provides powerful solutions to process huge data sets on different types of data processors from statistical analysis to machine and deep learning. The disadvantages of such systems

The second type targeted to personificated medicine. This solution presented a list of tools for collecting and systematization patient data. It can help in relative process medical data of a patient and notify therapist about changes in health state of the patient. One of the disadvantages such system that it works only with patient medical card or manual entered by patient information. One of the examples of this system is IBM Watson Care Manager [8].

The disadvantages of the first type of systems that they can not be used for personification medicine. They can be only a part of it with help to increase the speed of research and creating top level of prediction and decision system in healthcare. The second type looks more like user data aggregator to help the personal doctor better understand changes in patient health and real-time monitoring it. We think that combination of both types can fully show the potential of using IoT and Machine Learning in personificated medicine.

1.3 Objectives, Approach and Structure

One of the biggest question in creating big decision support system for healthcare is selecting environment for its rollout. Best and more secure way it set up all system by yourself. But on design stage better is focus on the application, then on preparing hardware, setting the private cloud, thinking about server stability and specific hardware for running ML. The best way is using cloud services for it. To better understanding of cloud service system requirements was proposed to create a small prototype of all system to understand, how we can connect medical devices to the cloud, how much data they can generate, what type of hardware we need to run ML for data processing.

As an example, we use ECG-device for generating data and real-time transferring data to the cloud. All collected medical data is storing in the database. Then data is processing with ML and results are storing on the next level of the database. This example helps to understand the amount of data what system can receive, and what computing power needed to process it.

The remainder of the paper is structured as follows. Section 2 presents a brief description of application architecture. Section 3 describes simple medical sensor and type of data generated by sensor and type. Section 5 presents a comparison of cloud services followed by concluding remarks.

2 Application Architecture

The application should contain few different levels each of them must produce, store or process different types of medical data. All levels are separated from each other to satisfy single responsibility principle.

Level 1. Hardware. This level includes all devices (sensors) that provide medical data about the patient. It can be EKG sensors, thermal sensors, physical activity sensors, tonometers, blood glucose monitoring devices and etc.

They collect data and send it to the first database level, in the cloud. These do not do any preprocessing data. They can send data directly to the cloud with WI-FI or GSM, or use some Gateway [9]. To Gateway they can connect with Bluetooth, ZigBee or wired connection. Example of such architecture is depicted in Fig 4.

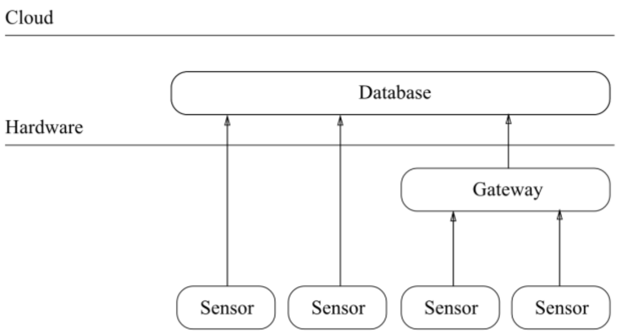

Fig. 1. Communication of hardware and other levels.

Level 2. Database. This is the first level that store all raw data from sensors. Al-so, on this level, it is possible to connect external medical databases. The external medical database may provide raw data of medical survey form different medical insti-tutions.

Also, on this level can be connected free databases with datasets of medical survey for training artificial intelligence (AI).

Level 3. First Level of Machine Learning. On this level, different AI for pro-cessing medical datasets are kept. AI on this level is very simple and can detect only one type of heels problem. As an example, it can be few AI for processing ECG, EEG data, ultrasound survey, fluorography and magnetic resonance imaging (MRI image processing). All processed data are stored on the second level of the database.

Level 4. Second Level of Database. This level stores all processed data of pa-tients from different sources and after AI algorithms. Also, on this level more infor-mation about patients like patient history and patient circumstances is kept. This level collects all related to patient data for next step of processing.

Level 5. Prediction of AI. There is a big and complex AI for prediction the pa-tient health state on this level. This is the most complex AI in whole system because it works with a big amount of different data. As an example, on this level for predic-tion heart disease can be used not only ECG of the patient. Age, gender, information about similar problems from parents and grandparents, information about physical activity on time getting ECG are taking into account on this level. Even it can take into account information about the ecological situation in place where the patient is living.

All levels of healthcare IoT application are shown in Fig. 2.

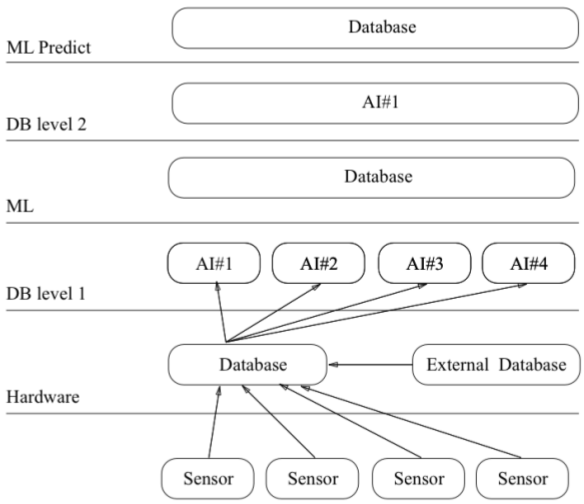

Fig. 2. All healthcare IoT application levels.

3 Collecting and Processing Medical Data

As an example, to calculate how many data system can receive, Heart Rate Monitor based on chip AD8232 [10] was used. It produces eighty states electrical activity of the heart over one second. All produced data collecting during some period (from 15 sec. to 48 hours [11]). All this information is stored in a special file. A full timeline of measurement, short patient information (id of the patient, id of the survey, type of measurement, time, duration) are stored in the small blocks of data and then sent to the cloud.

To automate the process of classification of medical data was proposed to use neural networks (NN). The simplest NN for ECG signals classification is based on the statistical algorithm called KernelFisher Discriminant analysis [12]. In a NN, the operation is organized into a multi-layered feed-forward neural network [13] with three layers namely Input layer, hidden layer and Decision layer. Such neural network can process raw data from seasons and make classification of them with mostly human accuracy.

4 Comparison of Cloud Services

Nowadays there is a wide specter of the cloud services that can be used for any type of projects. Selected cloud service should satisfy a few basic demands in case of deploying AI application:

- support last types of NVidia GPUs;
- big amount of memory (more than 64GB RAM and 16 GB video RAM);
- native support of different types of machine learning engines;
- interactive tools for data exploration, analysis, visualization and machine learning.

Creating Healthcare AI application imposes few additional demands as:

- reliability;
- support HIPAA;
- safety of patient data

Based on requirements above three cloud services was compared: Google, IBM, Amazon.

IBM. IBM Watson Machine Learning [14] is a full-service Bluemix offering that makes it easy for developers and data scientists to work together to integrate predictive capabilities with their applications. The Machine Learning service is a set of REST APIs that make integration with programming language to develop applications that make smarter decisions, solve tough problems, and improve user outcomes.

Watson Machine Learning allows to create different models and compare the results. Create automated experiments and self-learning models. Also, Watson Machine Learning allows easy visualize created models. It allows creating machine learning models using visual modeling tools and quickly identify patterns, gain insights, and making decisions faster. That is important on the first steps since it allows to involve doctors without programming knowledge in the first stage of creating predictions models.

Also deploying application on IBM Bluemix bring availability to connect IBM Watson Health. It brings access to medical databases for training models and improve machine learning models.

Google AI. Google Cloud Machine Learning Engine [15] is a managed service that enables to easily build machine learning models that work on any type of data of any size. The service is integrated with Google Cloud Dataflow for pre-processing, allowing to access data from Google Cloud Storage, Google BigQuery, and others.

One of the advantage and disadvantage of using Google Cloud AI is supported only TensorFlow SDK for machine learning. TensorFlow one is the most popular now frameworks for machine learning which allow works with different types of data such as simple data (sets of some calculations or datasets), processing images, video, and audio.

Working with TensorFlow give us the ability to create Portable Models. It can be used the open source TensorFlow SDK to train models locally on sample data sets and use the Google Cloud Platform for training at scale. Models trained using Cloud Machine Learning Engine can be downloaded for local execution or mobile integration.

Amazon. Amazon like previous companies provides similar scope of services for machine learning. AWS Machine Learning [16] allow creating big scalable machine learning system based on all the major frameworks, including TensorFlow, Caffe2, and Apache MXNe. One of the advantages is not only CPU and GPU for data processing.

AWS allows to create optimized instances based on combination CPU, GPU and even FPGAs with allow get same performance with less cost.

It can be noticed that all giants of internet commerce provide mostly similar tools for creating machine learning systems.

Many of they provide only special type of cloud which allows to create scalable ML systems. All 'containers' have access to powerful CPU and GPU, so selecting of cloud depends on used framework for machine learning, price, and infrastructure of whole project.

All information about supported frameworks is shown in Table 1.

Table 1. Supported Machine Learning libraries.

	TensorFlow	Caffe	Caffe2	Apache MXNet	Keras	Gluon
Google Cloud AI	x					
AWS Machine Learning	x	x	x	x	x	x

IBM Watson	x	x	x	x	x

Processing of data with machine learning require huge computation powers. Using specialised systems decrease cost of all system and increase it speed. GPUs in this case one of the better decision because they have more computational units and having a higher bandwidth to retrieve from memory. Now the most popular GPU for machine learning is NVIDIA TESLA P100 and NVIDIA TESLA K80. All configurations of instances displayed in Table 2 and Table 3.

Table 2. Min and Max configuration of virtual PC based on NVIDIA TESLA P100.

	CPUs	GPUs	GPU RAM	RAM
Google Cloud AI	1-64 cores	1-4	16-64 GB HBM2	1-208 GB
AWS Machine Learning	1-64 cores	1-4	16-128 GB HBM2	61-732 GB
IBM Watson	12 cores	1	16 GB HBM2	64 GB

Table 3. Min and Max configuration of virtual PC based on NVIDIA TESLA K80.

	CPUs	GPUs	GPU RAM	RAM
Google Cloud AI	1-64 cores	1-8	12-96 GB HBM2	1-208 GB
AWS Machine Learning	4-64 cores	1-16	12-192 GB HBM2	61-732 GB
IBM Watson	16-24 cores	2	16 GB HBM2	64 GB

After comparing of the cloud services, it can be finally analyzed which service is better for deploying healthcare IoT application. Google is one of the best decision for TensorFlow. It provides powerful instances for running machine learning, easy integration to others Google services like Big Query and others. Also, for the lightweight solution on small IoT devices like Raspberry-Pi or android smartphone, it provides tools for porting trained models to TensorFlow-Lite. In its turn, IBM is the best solution for integration with other IBM Bluemix services and IBM Watson Health. It supports many different frameworks for ML and powerful hardware for it. Also, the big plus is access to medical datasets of American hospitals for training [17]. Amazon is the best solution to deploying AI application. It provides most powerful and scalable hardware for ML including CPU, GPU and FPGA [18], supports the widest set of ML libraries and services for storing and processing data. Also, there are many libraries and integrations of hardware (sensors and gateways) directly to Amazon services.

5 Conclusions and Future Work

This paper describes the first step in creating IoT Healthcare cloud service. The first prototype of application architecture and data flow from hardware to predictions and its results were described. According to the application architecture and data sources was made comparing cloud services and selected better variant for prototype deploy.

Next step will be dedicated to increasing count and types of medical sensors. Add additional sources of data like meteorological reports. Move from Prototype to first MVP project where we can start to test this solutions on real peoples. In addition, it is planned to focus the solution to collecting and processing data from handheld devices and smartphones and smartwatches.

Acknowledgements

This paper results from the Erasmus+ programme educational project ALIOT «Internet of Things: Emerging Curriculum for Industry and Human Applications» (reference number 573818-EPP-1-2016-1-UK-EPPKA2-CBHE-JP, web-site http://aliot.eu.org) in which the appropriate course is developed (ITM4 - IoT for health systems) within its framework, we have developed modules related to IoT systems modelling. The authors would like to thank colleagues on this project, within the framework of which the results of this paper were discussed. The authors also would like to show their deep gratitude to colleagues from Computer systems, networks and cybersecurity, National Aerospace University «KhAI» for their patient guidance, enthusiastic encouragement and useful critiques of this paper.

References

1. Strielkina A., Uzun D., Kharchenko V.: Modelling of healthcare IoT using the queueing theory. Intelligent Data Acquisition and Advanced Computing Systems (2017). doi: 10.1109/IDAACS.2017.8095207
2. Internet of Things - number of connected devices worldwide 2015-2025 (2018): https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/
3. Ericson Mobility Report (2012): https://www.ericsson.com/assets/local/news/2016/03/ericsson-mobility-report-nov-2015.pdf
4. Nick Earle: 50 Billion Things, Coming to a Cloud Near You (2017) https://blogs.cisco.com /news/50-billion-things-coming-to-a-cloud-near-you
5. Kenji Suzuki: Pixel-Based Machine Learning in Medical Imaging. International Journal of Biomedical Imaging (2012). doi:10.1155/2012/792079
6. Muhammad Wildan Gifari, Hasballah Zakaria, Richard Mengko: Design of ECG Homecare:12-lead ECG acquisition using single channel ECG device developed on AD8232 analog front end (2015). doi: 10.1109/ICEEI.2015.7352529

7. Google, Healthcare & Life Sciences Solutions: https://cloud.google.com/solutions/healthcare-life-sciences/
8. IBM, Watson Care Manager: https://www.ibm.com/us-en/marketplace/watson-care-manager
9. Introductory guide to identifying ecg irregularities. http://www.dcbiomed.com/ proimages/materials/ [accessed December 20, 2017]
10. Yongjin Kwon, Kyuchang Kang, Changseok Bae: Unsupervised learning for human activity recognition using smartphone sensors: Expert Systems with Applications pp. 6067-6974. doi: 10.1016/j.eswa.2014.04.037
11. Pravin Kshirsagar and Sudhir Akojwar,“Hybrid Heuristic Optimization for Benchmark Datasets”, International Journal of Computer Application (0975-8887), Vol.146-No.7, July 2016.
12. Pravin Kshirsagar and Sudhir G. Akojwar,”Prediction of Neurological Disorders using Optimized Neural Network” International conference on Signal Processing, Communication, Power and Embedded System (SCOPES),Oct. 2016 .
13. Pravin Kshirsagar ,Vijetalaxmi Pai and Sudhir Akojwar ” Feature Extraction of EEG Signals using Wavelet and Principal Component analysis” National Conference On Research Trends In Electronics, Computer Science & Information Technology And Doctoral Research Meet, Feb 21st & 22nd.
14. Watson Machine Learning: https://www.ibm.com/cloud/machine-learning [accessed January 10, 2018]
15. Cloud ML Engine Overview: https://cloud.google.com/ml-engine/docs/technical-overview [accessed January 11, 2018]
16. Machine Learning on AWS: https://aws.amazon.com/machine-learning/ [accessed January 13, 2018]
17. Healthcare Analytics – IBM Watson Health: https://www.ibm.com/watson/health/explorys/healthcare-analytics/ [accessed January 9, 2018]
18. El Naqa, Ten Haken: Quantum-based machine learning for oncology treatment. Journal of Computer Science & Systems Biology (2017). doi: 10.4172/jcsb.1000143