Architecture for collecting and analysing data from sensor devices

Dona Jankova[1], Ivona Andova[1], Merxhan Bajrami[1], Martin Vrangalovski[1], Bojan Ilijoski[2], Petre Lameski[2] and Katarina Trojachanec Dineva[2]

[1] LOKA, 350 2nd Street, Suite 8 Los Altos, CA 94022, USA
[2] Ss. Cyril and Methodius University, Skopje, North Macedonia
`dona, ivona, merdzhan, martin@loka.com,`
`bojan.ilijoski, petre.lameski, katarina.trojacanec@finki.ukim.mk`

Abstract. A sensor device measures one or many health indicators for the user wearing it. The vital parameters are monitored and displayed to an individual or to a medical person in real or orchestrated time. This paper focuses on creating a general architecture that enables collecting, analyzing and transmitting data depending on the user needs. It can be used to track changes in vital parameters, without the need of the individual being physically present. With the focus on smartwatch data, as an example of a wearable device, this paper shows the possibilities for its application in finding accuracy of true authentic emotions that a person with autism spectrum disorder is feeling, by measuring the person's vital parameters.

Keywords: wearable devices · sensors · application · data · infrastructure · framework.

1 Introduction

In the past few years, a rapid development of wearable technology devices and their application in many areas, especially the healthcare industry, was witnessed [28]. With the intense development of information and communication technologies, the wearable technology ignited a new way of human-computer interaction [20, 11]. A wearable device is defined as a smart electronic device, small enough to be worn as an accessory, embedded in clothing or implanted in the user's body, with the ability to measure, collect, display and even transmit health data in real time, through sensor integration [21, 32]. One example is for medical purposes, where the physician would use this data to track the consumer's vital signs [13, 4] such as: temperature, blood pressure, blood oxygen, heart rate, physical movement and electrodermal activity of the heart, muscles, brain and skin. This paper proposes an application architecture that can work for any sensor measuring biomedical parameters, regardless of whether it is wearable or not, like in the case of [2] for analyzing volatile organic compounds (VOCs) in the exhaled breath. The architecture can ingest the sensor data, preprocess them and output some results to the user. In the diagram shown below (see Figure 1), the candidate represents the person from whom the data are being collected. A sensor device can be: a camera that inspects and observes the person's movements, an audiometer that measures the sound decibels, smart glasses that give information about vision and motions etc. All of these can help people with diseases such as in the case of Parkison's [31].

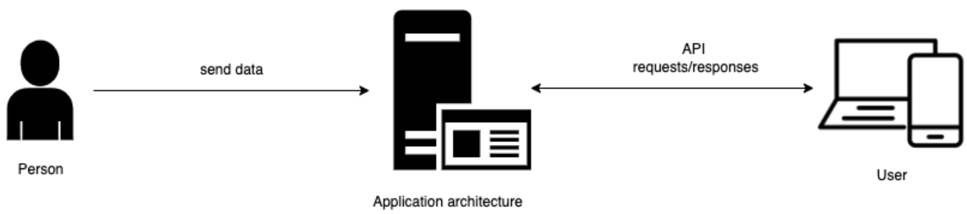

Fig. 1. High level approach on architecture diagram

After the application starts ingesting real time data, it can preprocess it, store or send the results to the client regarding the application implemented logic or user needs (see Figure 1 - API requests/responses). Another thing that this architecture can implement is alerting someone when a certain threshold is reached or some requirements are met [18].

2 Related work

One way of reshaping the future is using robots in everyday life. Paper [30] proposes a socially adaptable framework for human-robot interaction examining the interaction between a human and a robot in the role of a small child. Each session consists of two parts: one is when the robot can move, and the other part is when the robot can only sit and move its arms and head. The human goal is to animate the robot and examine how often the robot reaches the limit of animation saturation. In paper [1] a novel mathematical model has been proposed based on an adaptive multi-robot therapy of ASD children. The proposed model uses a multi-robot system as therapist, without any external stimuli (from environment), to improve the skills of the autistic child. Two methods of robot interaction were conducted: with and without interaction between the robot and the child. The idea of this paper was to design and develop a single mathematical model for adaptive multi-robot based therapy. The advantage of this model is that it does not require continuous engagement of the human therapist.

Another way is to improve the utilization of wearable device data using resources and services from Amazon Web Services (AWS) by creating an architecture as proposed in article [10]. It can be used to help people with autism and their caregivers as described in paper [25], by monitoring data from multiple signals and informing the caregiver of their present feeling.

Looking from a different perspective, creating a mobile application can be very useful and also a fun way to help people with autism. It is described in the papers [17, 15, 14, 16]. Paper [17] focuses on creating a mobile charades-style game, Guess What?. This game challenges the child to act out a series of prompts displayed on the smartphone held on the care provider's forehead and the parent attempts to guess what emotion the child is acting out (surprised, scared, or disgusted). During this time the application collects and processes a video recording of the child in order to detect emotion. This serves as an indicator to what the child likes and feels more comfortable doing. Providing real-time feedback and adapting game difficulty in response to the child's performance necessitates the integration of emotion classifiers into the system. Due to the limited performance of existing emotion recognition platforms for children with ASD, paper [16] proposed a novel technique to automatically extract emotion-labeled frames from video acquired from game sessions, which the authors hypothesize can be used to train new emotion classifiers to overcome these limitations. In paper [15] labeling those images is better described using several algorithms for extracting semi-labeled frames from these videos. With the small dataset gathered from this game, emotion classifiers are tested and described in the paper [14]. Emotion classifiers available off-the-shelf to the general public through Microsoft, Amazon, Google, and Sighthound, are well-suited to the pediatric population, and could be used for developing mobile therapies. This study aimed to test these classifiers. The findings suggest that commercial emotion classifiers may be insufficiently trained for use in digital approaches to autism treatment. The idea is that once additional video data are available, the methods described in paper [15] will be employed to generate a large labeled dataset that will be used to train convolutional neural network classifiers for emotion recognition that are robust across differences in age and developmental delay.

When dealing with emotions, paper [22] presents novel tools for analysis of human behavior data regarding robot-assisted special education. The objectives include, first, an understanding of human behavior in response to an array of robot actions and, second, an improved intervention design based on suitable mathematical instruments. To achieve these objectives, Lattice Computing (LC) models in conjunction with machine learning techniques have been employed to construct a representation of a child's behavioral state. Using data collected during real-world robot-assisted interventions with children diagnosed with Autism Spectrum Disorder (ASD) and the aforementioned behavioral state representation, time series of behavioral states were constructed. Four computational tools were developed for behavioral data analysis including (1) histograms of states per robot modality, (2) histograms of state transitions per robot modality, (3) a visualization of time series of states and (4) histograms of robot actions per robot modality. The idea of this proposed behavioral analysis can allow psychologists to design robot-assisted interventions. From the analysis presented here, which involves all cumulatively gathered data for all children across all activities, animations and the use of LEDs appear to prompt children to turn their heads toward the robot more frequently than other modalities, which suggests an increase in engagement.

For physically demanding industries sensor integrated devices can be used for monitoring biomedical changes of workers' behaviour and health, such as in the case of construction workers [26, 8].

3 Architecture

3.1 Collecting data

The architecture flow begins with collecting data. For that purpose a sensor device connected to the internet is needed in order to send the data to some kind of end point or store them in a database. One device can have multiple sensors [9, 5, 27]. With the aim of utilizing the architecture for monitoring a medical disorder such a ASD, a wearable device was put on a participant's wrist with 4 sensors: accelerometer (measures changes in speed of movements), photoplethysmograph (measures blood volume pulse - BVP), electrodermal activity of the skin and temperature sensor. Additionally from the BVP sensor another measure was extracted, that is the heart rate. All this data is sampled from the sensors at different frequencies. Every 2 seconds, all the data collected during that interval is sent to an end point.

3.2 Data preprocessing

All the data that is collected is some kind of time series data, which means that in order to use it, they need to be preprocessed [19]. First, if there are any faulty values, they are removed. Then there is a need for matching the sampled frequencies of the different sensors in the same device. Each sensor is sampled at a different interval, therefore the data are needed to be duplicated in order to match the highest sample frequency. After that was sorted out, a statistical analysis was needed in order to create features that fit the previously created model's input. In this case, every attribute represents some kind of statistical analysis of the sensor data during that 2-second interval. It includes statistics such as min, max, average, median absolute deviation etc.

3.3 Analyzing data and ML models

After creating the features, they are fed into a model that was pre-trained on the same kind of sensor data. This can work on a wide range of different models such as: classification, clustering, regression etc., depending on the specific user needs, in order to derive an conclusion. Here, a classification approach was implemented, having an emotion as an output (see Figure 2) [29].

Based on the performances, model stacking approach was chosen as the best one: ModelStacking(L0: Knn/SVM/LDA/RandomFores/XGBoost/dTree/AdaBoost, L1: Xgboost) with accuracy 0.82, precision 0.78 and recall 0.71.

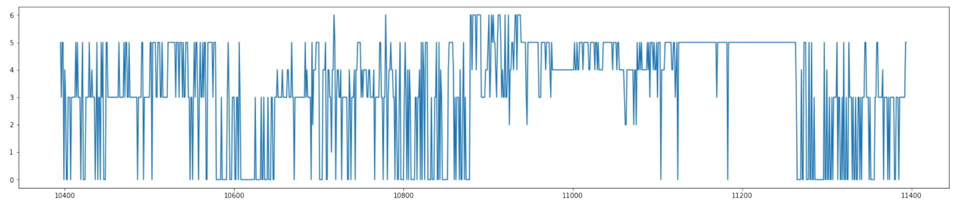

Fig. 2. Model output distribution visualization

3.4 Visualization

The results can be stored in another database on the cloud. Besides that, they can be streamed to a client host, depending on the needs. It can be a Web application or a mobile application that displays the results. Also, with the usage of Lambda functions, processed data can also be displayed in order to show valuable information to the user. For example Figure 3 shows average sensor data and results from a neural network model that predicts emotions from the sensor data.

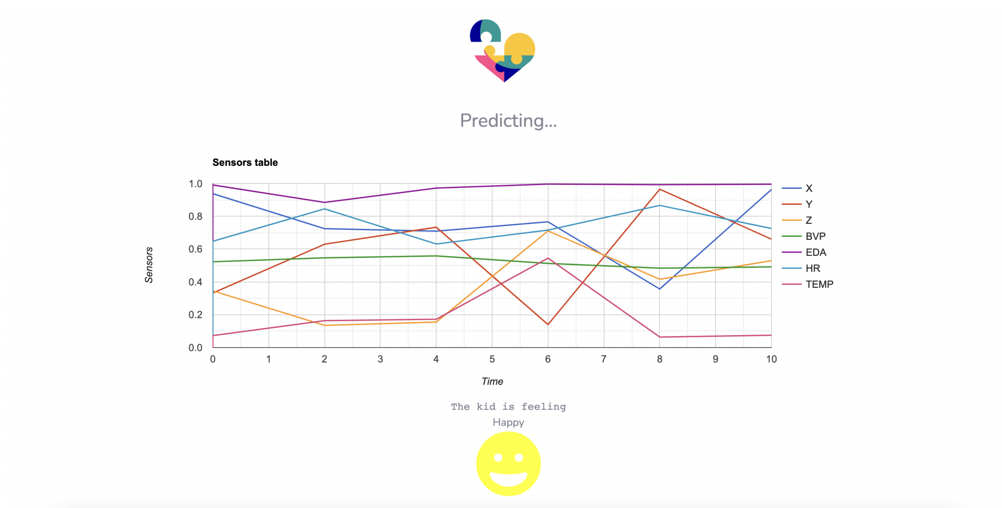

Fig. 3. Data visualization and model output

3.5 Application flow

Application architecture implemented on Amazon Web Services (AWS)
This paper shows an architecture implemented on Amazon Web Services. Although, the same can be done in any other cloud computing platforms such as:

Google Cloud Platform and Microsoft Azure. This implementation shows how any wearable device can be used as a source of ingesting data.

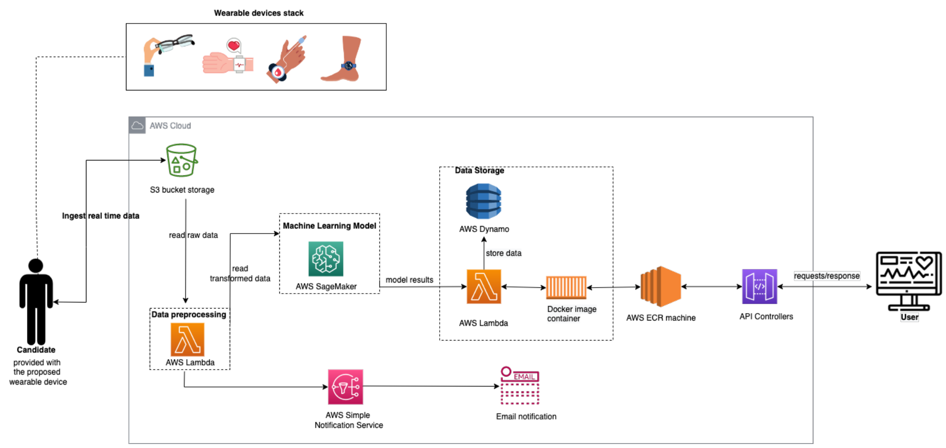

Fig. 4. Architecture flow implemented on AWS

The architecture diagram on Figure 4 shows a high-level approach of implementing the infrastructure for the wearable devices. There is a wearable devices stack that consists of multiple wearable devices, where each of them is similar to the others and can be easily replaced depending on the user's needs. The application user has to establish a connection with the sensor device. Once the application user (parent, doctor, trainer, supervisor etc) has successfully established the connection, the sensor will be sending real time data. The data will be handled by AWS services: S3 bucket for storing the data, Lambda for data preprocessing, Sagemaker for creating machine learning models used for prediction. After the model outputs the results, the application user has the option to choose how to handle the data: store it in AWS DynamoDB database or display it on the monitor.

AWS Services explanation The following section shows all the resources and that were used for the whole process in the architecture flow:

Amazon Simple Storage Service (Amazon S3) [12]is a scalable object storage service where the raw data files from the sensors are put.

AWS Lambda in this case has multiple purposes. It is used as a compute service that preprocesses the data, in order to transform them for model input. It also reads the model output and can store it into the database or stream the data depending on user needs. Another function Lambda does is triggering the docker image building process.

Amazon DynamoDB is a NoSql database that delivers single-digit millisecond performance at any scale and is used to avoid processing of duplicate files. This database is chosen because it provides high throughput at a very low latency which is very important for reading data if the application user chooses to display the data.

Amazon Simple Notification Service (Amazon SNS) is a fully managed messaging service for both application-to-application (A2A) and application-to-person (A2P) communication that is used to send alerts to the application user for the connection status of the wearable device. It will inform whether the connection is established or not.

AWS Cloud Control API is a set of common application programming interfaces (APIs) that make it easy for developers and partners to manage the lifecycle of AWS and third-party services. It is responsible for handling API calls, such as requests and providing the application with responses.

AWS SageMaker is a cloudformation tool for machine learning purposes, used for building and training machine learning models. Depending on the project purpose multiple models can be trained on the Sagemaker instance or the instance itself can be used for statistical data analysis that can be provided to the application user.

A Docker [7, 24] image is a file used to execute code in a Docker container. Docker images act as a set of instructions to build a Docker container, like a template. It contains application code, libraries, tools, dependencies and other files needed to make an application run

AWS EC2 machines [12, 23, 3] provide scalable computing capacity in the Amazon Web Services (AWS) Cloud. An EC2 instance in the architecture will be used for building and running the docker image.

Architecture of the web application used for children's with autism One way of using this architecture is in a web application developed for transmitting health data in real time from children with autism spectrum disorder to medical professionals, enabling health tracking without the individual being physically present [6]. People with autism are obstructed to understand, process and communicate emotions to other people, which explains the importance of emotional vital signs monitoring, in order to assess the individual's well-being. The data is being collected through a smartwatch containing sensors as an example of a wearable device throughout a therapy session with a doctor. Using a machine learning trained model the application can predict how the candidate is feeling during the session. The application is built and deployed on AWS and is used for visualizing the data and showing results of the model to medical personnel regarding the candidate's feelings (see Figure 5).

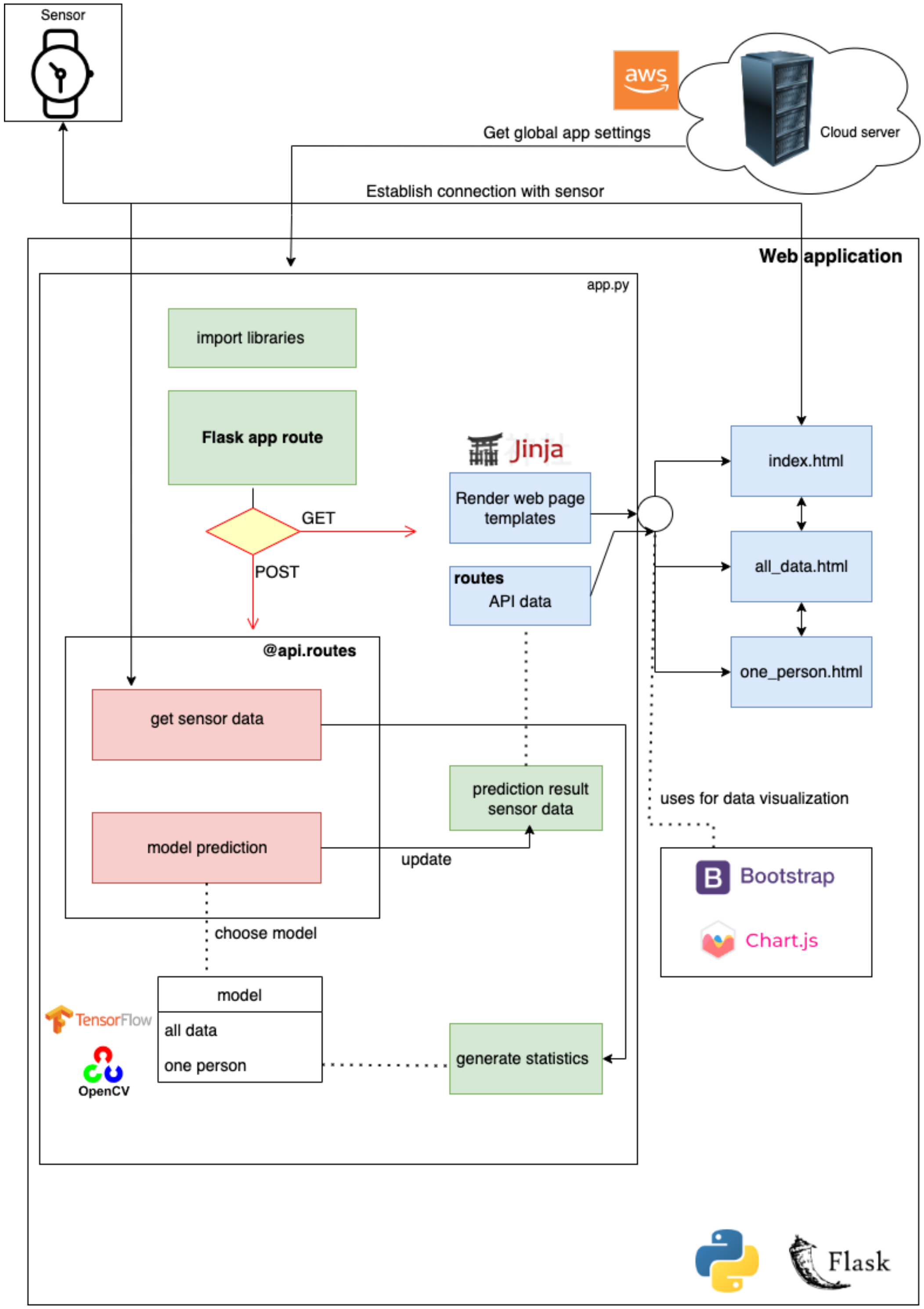

Fig. 5. Web application architecture

The RESTful application is built with the Flask framework. Once the doctor runs the application, one will have the option to establish a device connection. Since every sensor will be generating raw data at a different frequency, data pre-processing is needed in order to match each sensor to the one with highest frequency. Every 2 seconds the data collected during that interval are being sent and handled from an API gateway in real time, where via POST routes the application will analyze the data. For that a script will generate statistics (features) creating the data format that fits the model input. Next, the doctor can choose a model approach that will be used for predicting emotions. Once the doctor chooses it, the application will render a page where the preprocessed data and the model output will be visualized. Google Chart.js is used for data visualization and Bootstrap framework for styling. For rendering the templates via GET routes of the API, Jinja is used. The front pages that the doctor will control will be visualizing the data in real time and displaying the output model prediction.

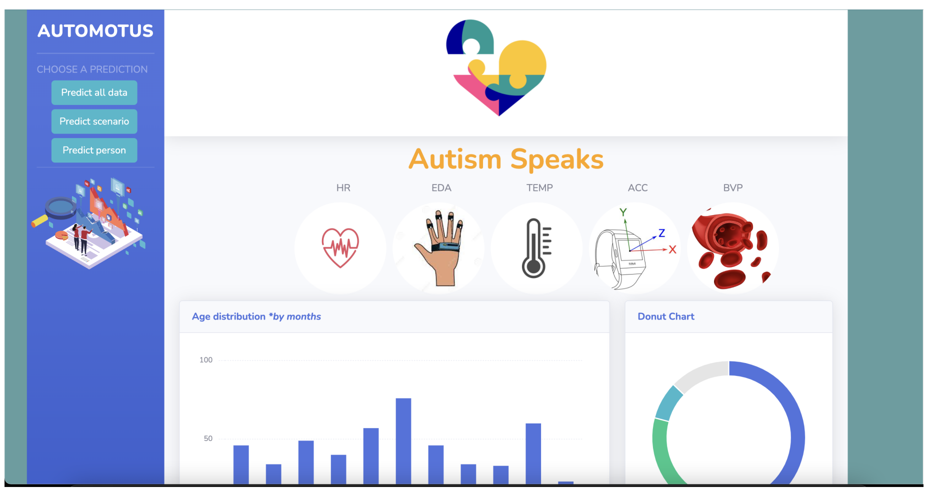

Fig. 6. Application UI for home page

4 Conclusion

What does it take to form the new tomorrow and how do these novelties blend into what is known today? The answer is through changes. Change is a necessity. Change is a vital process, in order to build a better tomorrow. Because the increased sensor installations into objects in the human surroundings experience an exponential growth and sensors are starting to get embedded in all kinds of devices, it is essential to make a greater use out of their data. Hence, this paper focuses on providing a platform for collecting data from sensor devices (both wearable and not), making them meaningful. Data can be interpreted differently, depending on the application's context, therefore the architecture is not limited to the domain usability or appliance. The concept is the same: architecture that collects and transmits health data from sensor devices in real time and displays it to a person of interest (the user using the device or another persona - depending on the specific logic/need), so an interpretation can be made. The proposed architecture covers all the challenges related to sensor utilization, and its flexible and adjustable concept can solve any challenges with little or no adaptations. For example, If the user has a machine learning model and wants to view patterns in customer's behavior, they can make their model a central part in the architecture, and apply it to practice in order to address their operational needs and interpret their data collection. Or if a service is not appropriate for meeting the operational goals, it can be easily swapped. Simple as that, with the benefits of cloud computing. This described architecture solves the latency issues associated with data processing that can occur when sending data in real time, shaping the future in a way that transmitting and processing data in real time, without a delay, in order to make a data collection significant, is the new normal. The cloud computing concept will reduce the IT operating costs in a way that the user will get its work done, without the need of new hardware, expenses or inconveniences. Depending on the needs, it will provide effective up scaling or down scaling on the resources. Since security is a concern when working with data, Amazon Web Services ensures data confidentiality and prevents data leakage, through its available services. The proposed architecture finds use in tracking users regardless of their location, thus the gathered data can give insights, without the individuals being physically present. In view of the fact that the architecture can easily adapt to the area of interest, one adaptation was in the field of autism spectrum disorder. The architecture addressed the challenges related to gathering, transmitting and analyzing smartwatch data from children with autism, that lead to the conclusion that the authenticity of the emotions can be and is related to the data gathered from the sensors. Therefore, for this case, the data can be used for understanding how people truly feel, regardless of the emotion they are expressing on the outside.

Acknowledgement

The work in this paper was partially financed by the Faculty of Computer Science and Engineering, Ss. Cyril and Methodius University in Skopje.

References

1. Ali, S., Mehmood, F., Dancey, D., Ayaz, Y., Khan, M.J., Naseer, N., Amadeu, R.D.C., Sadia, H., Nawaz, R.: An adaptive multi-robot therapy for improving joint attention and imitation of asd children. IEEE Access **7**, 81808–81825 (2019). https://doi.org/10.1109/ACCESS.2019.2923678

2. Alkhouri, N., Singh, T., Alsabbagh, E., Guirguis, J., Chami, T., Hanouneh, I., Grove, D., Lopez, R., Dweik, R.: Isoprene in the exhaled breath is a novel biomarker for advanced fibrosis in patients with chronic liver disease: A pilot study. Clinical and translational gastroenterology **6**(9), e112 (2015)
3. Bhise, V.K., Mali, A.S.: Cloud resource provisioning for amazon ec2. In: 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT). pp. 1–7. IEEE (2013)
4. Boudargham, N., Abdo, J.B., Demerjian, J., Guyeux, C.: Exhaustive study on medical sensors. In: International Conference on Sensor Technologies and Applications (2017)
5. Brida, P., Krejcar, O., Selamat, A., Kertesz, A.: Smart sensor technologies for iot (2021)
6. Cabibihan, J.J., Javed, H., Aldosari, M., Frazier, T.W., Elbashir, H.: Sensing technologies for autism spectrum disorder screening and intervention. Sensors **17**(1), 46 (2016)
7. Cito, J., Ferme, V., Gall, H.C.: Using docker containers to improve reproducibility in software and web engineering research. In: International Conference on Web Engineering. pp. 609–612. Springer (2016)
8. Edirisinghe, R.: Digital skin of the construction site: Smart sensor technologies towards the future smart construction site. Engineering, Construction and Architectural Management **26**(2), 184–223 (Jan 2019). https://doi.org/10.1108/ECAM-04-2017-0066, https://doi.org/10.1108/ECAM-04-2017-0066
9. Elayan, H., Shubair, R.M., Kiourti, A.: Wireless sensors for medical applications: Current status and future challenges. In: 2017 11th European conference on antennas and propagation (EUCAP). pp. 2478–2482. IEEE (2017)
10. Engdahl, S.: Blogs (2008), https://aws.amazon.com/blogs/industries/improving-the-utilization-of-wearable-device-data-using-an-aws-data-lake/
11. Ernst, T., Guillemaud, R., Mailley, P., Polizzi, J., Koenig, A., Boisseau, S., Pauliac-Vaujour, E., Plantier, C., Delapierre, G., Saoutieff, E., et al.: Sensors and related devices for iot, medicine and s mart-living. In: 2018 IEEE Symposium on VLSI Technology. pp. 35–36. IEEE (2018)
12. Garfinkel, S.: An evaluation of amazon's grid computing services: Ec2, s3, and sqs (2007)
13. Howard, R.M., Conway, R., Harrison, A.J.: A survey of sensor devices: use in sports biomechanics. Sports biomechanics **15**(4), 450–461 (2016)
14. Kalantarian, H., Jedoui, K., Dunlap, K., Schwartz, J., Washington, P., Husic, A., Tariq, Q., Ning, M., Kline, A., Wall, D.P., et al.: The performance of emotion classifiers for children with parent-reported autism: quantitative feasibility study. JMIR mental health **7**(4), e13174 (2020)
15. Kalantarian, H., Jedoui, K., Washington, P., Tariq, Q., Dunlap, K., Schwartz, J., Wall, D.P.: Labeling images with facial emotion and the potential for pediatric healthcare. Artificial intelligence in medicine **98**, 77–86 (2019)
16. Kalantarian, H., Jedoui, K., Washington, P., Wall, D.P.: A mobile game for automatic emotion-labeling of images. IEEE transactions on games **12**(2), 213–218 (2018)
17. Kalantarian, H., Washington, P., Schwartz, J., Daniels, J., Haber, N., Wall, D.P.: Guess what?: Towards understanding autism from structured video using facial affect. Journal of healthcare informatics research **3**, 43—66 (2019). https://doi.org/10.1007/s41666-018-0034-9, https://europepmc.org/articles/PMC7730314

18. Kostikis, N., Rigas, G., Konitsiotis, S., Fotiadis, D.I.: Configurable offline sensor placement identification for a medical device monitoring parkinson's disease. Sensors **21**(23), 7801 (2021)
19. Kotsiantis, S., Kanellopoulos, D., Pintelas, P., et al.: Handling imbalanced datasets: A review. GESTS international transactions on computer science and engineering **30**(1), 25–36 (2006)
20. Lee, J., Kim, D., Ryoo, H.Y., Shin, B.S.: Sustainable wearables: Wearable technology for enhancing the quality of human life. Sustainability **8**(5) (2016). https://doi.org/10.3390/su8050466, https://www.mdpi.com/2071-1050/8/5/466
21. Loncar-Turukalo, T., Zdravevski, E., da Silva, J.M., Chouvarda, I., Trajkovik, V., et al.: Literature on wearable technology for connected health: scoping review of research trends, advances, and barriers. Journal of medical Internet research **21**(9), e14017 (2019)
22. Lytridis, C., Kaburlasos, V.G., Bazinas, C., Papakostas, G.A., Sidiropoulos, G., Nikopoulou, V.A., Holeva, V., Papadopoulou, M., Evangeliou, A.: Behavioral data analysis of robot-assisted autism spectrum disorder (asd) interventions based on lattice computing techniques. Sensors **22**(2), 621 (2022)
23. Pham, T.P., Ristov, S., Fahringer, T.: Performance and behavior characterization of amazon ec2 spot instances. In: 2018 IEEE 11th International Conference on Cloud Computing (CLOUD). pp. 73–81. IEEE (2018)
24. Rad, B.B., Bhatti, H.J., Ahmadi, M.: An introduction to docker and analysis of its performance. International Journal of Computer Science and Network Security (IJCSNS) **17**(3), 228 (2017)
25. Ramasubramanian, K., Venkateswarlu, L., Lavanya, M.K., Unnati, L.: Emotional perception of individuals with autism spectrum disorder through machine learning and smart watch. Turkish Journal of Computer and Mathematics Education (TURCOMAT) **12**(13), 7217–7225 (2021)
26. Rao, A.S., Radanovic, M., Liu, Y., Hu, S., Fang, Y., Khoshelham, K., Palaniswami, M., Ngo, T.: Real-time monitoring of construction sites: Sensors, methods, and applications. Automation in Construction **136**, 104099 (2022)
27. Sh, I.: Ma, mahgoub i. Du E., Leavitt MA, Asghar W. Advances in healthcare wearable devices. Npj Flexible Electronics **5**(1), 1–14 (2021)
28. Shen, G.: Recent advances of flexible sensors for biomedical applications. Progress in Natural Science: Materials International (2021)
29. Siddiqui, U.A., Ullah, F., Iqbal, A., Khan, A., Ullah, R., Paracha, S., Shahzad, H., Kwak, K.S.: Wearable-sensors-based platform for gesture recognition of autism spectrum disorder children using machine learning algorithms. Sensors **21**(10), 3319 (2021)
30. Tanevska, A., Rea, F., Sandini, G., Cañamero, L., Sciutti, A.: A socially adaptable framework for human-robot interaction. Frontiers in Robotics and AI **7** (2020). https://doi.org/10.3389/frobt.2020.00121, https://www.frontiersin.org/articles/10.3389/frobt.2020.00121
31. Vera Anaya, D., Yuce, M.R.: Stretchable triboelectric sensor for measurement of the forearm muscles movements and fingers motion for parkinson's disease assessment and assisting technologies. Medical Devices & Sensors **4**(1), e10154 (2021)
32. Wanjari, N.D., Patil, S.C.: Wearable devices. In: 2016 IEEE International Conference on Advances in Electronics, Communication and Computer Technology (ICAECCT). pp. 287–290. IEEE (2016)