Ilkogretim Online - Elementary Education Online, 2021; Vol 20 (Issue 3): pp. 4124-4131
http://ilkogretim-online.org
doi: 10.17051/ilkonline.2021.03.425

An Unsupervised Cluster Analysis Framework For Wearable Iot Based On The Fog Computing

Vrince Vimal Department of Computer Science & Engineering, Graphic Era Hill University, Dehradun, Uttarakhand India, 248002 vvimal@gehu.ac.in

Surendra Shukla Department of Computer Science & Engineering, Graphic Era Deemed to be University, Dehradun, Uttarakhand India, 248002 surendrakshukla.cse@geu.ac.in

Bhasker Pant Department of Computer Science & Engineering, Graphic Era Deemed to be University, Dehradun, Uttarakhand India, 248002 pantbhasker2@gmail.com

Abstract

Wearable gadgets are quickly have become an increasing significant component of smart healthcare, which then in turn generates unique types of medical big data. These records are processed by the cloud and fog computing services, which is beneficial to medical care. The speedier data processing made possible by the Internet of Things (IoT) has been a benefit to the field of electronic healthcare. This work offers and assesses a strategy for carrying out machine learning with limited resources on Fog devices that are situated in close proximity to smartwatches for the purpose of incorporation into smart healthcare. In order to handle physiological data, the signals processing and computational modeling components of cutting-edge telecare systems have been moved to the cloud. This was done to facilitate remote access. In order to identify patterns in physiological data, we evaluated an initial form of an unsupervised classification algorithm large tool for data analysis that was created on top of a Fog platform. Our goal was to find these trends. Both the Intel Edison and the Raspberry Pi were employed as Fog computers in the proposed architecture. We used data on PD patients' home speech disorders acquired through routine monitoring to conduct our experiments (PD). Pathological speech data from smartwatch-wearing Parkinson's disease patients was analysed using the suggested architecture using machine learning. Results demonstrated the feasibility of the proposed architecture for clinical machine learning with modest hardware and software. It is possible that wearable IoT might be of service in smart telehealth settings if machine learning algorithms designed for the cloud could be adapted to work on edge computing devices like Fog Computing.

Keywords: Smart watch, Wearable, Telehealth, Signal processing, Machine learning algorithm

I. INTRODUCTION

Introducing Fog, an innovative design for computing, storing, management, and connectivity that places these functions in closer proximity to the end customers. To put it more plainly, this makes it possible for services to be moved to the "edge" of the networking. Fog computing allows computing and control functions to be inched closer to the sensor than cloud computing does, making fog the preferable choice. In order to understand speech signals for use in healthcare applications, the smart Fog architecture that we have presented makes use of the Fog idea. Speech digitization and machine learning are two techniques that can be utilised to recognize and monitor speech difficulties in people diagnosed with Parkinson's disease. This disease affects a sizeable portion of the world's population and is responsible for conditions such as dysarthria. It is possible for telemedicine monitoring in speech - language pathologists to be effective using innovative technology such as EchoWear. There are a number of indicators that show how dysarthria, speech prosody, and auditory characteristics are connected. Patients with Parkinson's disease almost usually experience dysarthria, according to the authors in . Having a breathy, loud voice, irregular pace, and slurred consonants; speech is monotonous with decreased stress. Extreme F0 variation and range was suggested by some authors who studied the speech of people with severe dysarthria. PD patients' speech has a high degree of multitudinousness, which is an important acoustic factor in dysarthria. The writers of a paper on Parkinson's disease and speech impairments call attention to a decrease in voice loudness, or hypokinetic dysarthria. Using unsupervised clustering, which we propose in this work as part of the SmartFog Fog Computing architecture, we are able to detect irregularities in PD patients' voice data (PD). At home, people with Parkinson's disease can monitor their progress with speech therapy by using a smartwatch. Information about the speaker's speech was sent to the Fog system via a nearby tablet or smartphone. The Fog machine records data on the loudness and fundamental frequency of aberrant speech. After establishing some sort of standard for the speech features, K-means clustering was used. Any time a noticeable change happens, the information is sent to the cloud for further examination. Sometimes the data is processed both locally and remotely. In this case, the Fog device can "intelligently" decide if and when to upload data to the cloud. Two different prototypes were developed, one using an Intel Edison and another using a Raspberry Pi. The two prototypes' processing speeds were compared.

Both approaches were tested using data gathered through telemonitoring of people with Parkinson's disease in order to determine how well they could spot irregular speech patterns. The proliferation of wearables in smart telehealth systems has resulted in an explosion of so-called medical big data. These files are utilized by specialists in the field of telemedicine in order to provide patients with superior care. Based on this research, it is suggested that low-resource machine learning should be deployed on Fog devices that are located close to wearables in order to promote smart telemedicine. Analyzing physiological data is done by modern telecare solutions using cloud-hosted signal processing and machine learning modules, much as it was done by traditional telecare systems, which used servers located on the premises to process the data. We carried out a K-means clustering analysis with the assistance of the typical fundamental frequency, measured in hertz (F0), as well as the typical intensity (in decibels). Using the fog platform, the software is able to successfully cluster the unlabeled data into groups of data that are comparable to one another. As one illustrative application, this method can be used in real time to categorise patients with Parkinson's disease into subgroups based on shared characteristics among their clusters.

II. LITERATURE SURVEY

With the Fog Architecture, data processing, transfer, and storage are all moved to the network's periphery. Several authors have each outlined their own take on Fog's architecture. The elements of FIT as outlined by **Monterio et al. (2016)** are as follows: The aforementioned smartwatch, fog computer, and cloud-based backend. In an effort to conceive WIoT in terms of its design, function, and applications, **Hiremath et al. (2014)** recently presented their findings. An example of a service-oriented design for fog computing may be seen in Dubey et al(2015) .'s presentation of The Fog Data. In this particular setting, the body of research emphasises the practicality and malleability of fog computing.

In order to make use of Big Data's recent advancements in the scientific and healthcare fields, tele-health has emerged. All sorts of informatics fields, including healthcare and medical computing, molecular bioinformatics, sensor information science, and others, stand to gain from individualised knowledge that has been gleaned from a diverse range of sources of information. Andreu et al. have confirmed and analysed the functionality of the Fog Information infrastructure (2015). A low-power embedded processor that is capable of performing data gathering and processing on material that has been received from a wide variety of sensing devices that are employed in hospital environments is the design that has been recommended. keeps bringing up the European initiative PERFORM, which develops a sophisticated multiparametric platform for the continued effective monitoring and evaluation of motor status in patients with Motor symptoms and other degenerative illnesses. It is a form of telemedicine that allows for remote monitoring of individuals who suffer from Parkinson's disease. The study provides an overview of the technological performance of the device as well as the statements made by patients regarding its application and the comfort it provided while it was being worn. In the course of our study, we analysed Parkinson's speech data and put it to the smart-fog architecture that we proposed. This technique strongly relies on Feature Extraction to function. Barik et al. (2018) use an optimised version of the K-means algorithm in order to cluster statistical properties such as the deviation of probability density functions of the retrieved features from the clusters. Cancela and colleagues have organised a library of feature vectors that were taken from Malay digits utterances and clustered them (2013). In this particular investigation, the returned characteristics.

III. PROPOSED METHODOLOGY

✓ Feature extraction

The very first stage of every machine learning study is feature engineering. Feature selection is the procedure of choosing the right data metrics to feed into a machine learning system. To help the K-means clustering analysis method discover the groups based on similarity, it is important to choose features that can capture the variability in the data. Subjects had Parkinson's disease, and we focused on two aspects of their speech: the fundamental frequency (F0) and the amplitude (Average). Parkinson's disease patients' speech patterns were recorded. Over the course of the analysis, 164 different snippets of conversation were taken into account. Praat is used as the scripting language for feature extraction . In order to determine pitch, the programme employs an accurate autocorrelation technique to identify acoustic periodicity. Sound intensity is determined by squaring sound value and then convolving with a Gaussian analysis window. Decibels are used to measure volume.

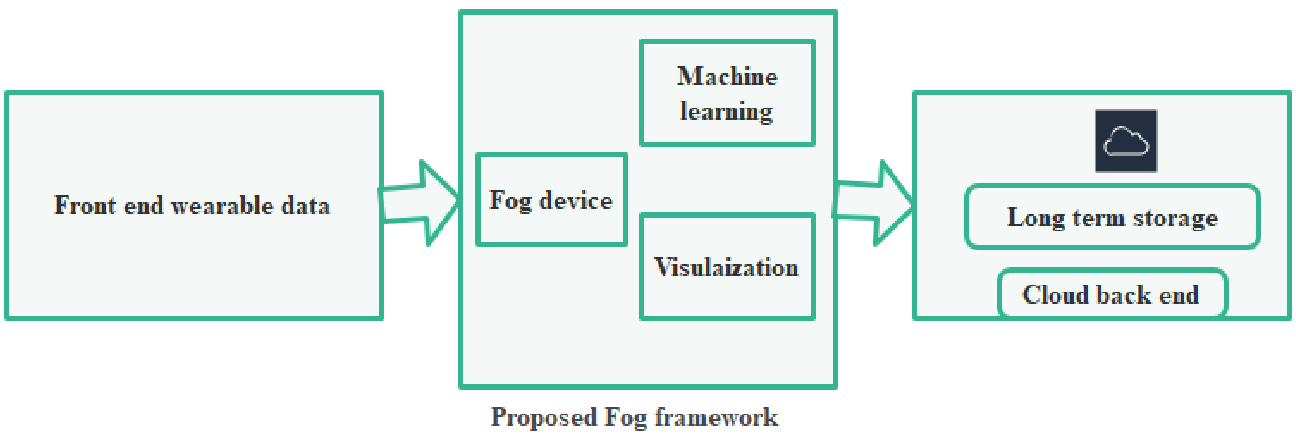

Fig 1: Smart Fog architecture for wearable Internet of Things analytics.

✓ K –means clustering

K-means clustering is a form of unlabeled data that can be utilized for exploratory data processing even in situations where there are no labelled inputs. The K-means algorithm is a wavelet packet method that finds widespread application in the field of data mining. Because the number of clusters, which will be referred to here as K, is of importance, the algorithm will concentrate on that particular feature of the material. Iterative manner placing each piece of evidence within one of the K groups that are indicated by the characteristics, the algorithm processes every point of data as it comes in. The characteristics and the K value are the two components that make up the algorithm's inputs. At the beginning of the process, K centroids are picked at random, and the algorithm then repeats until it reaches convergence. In order to achieve its goal, this method seeks to reduce the squared error function J.

$$j = \sum_{k=1}^{k} \sum_{i\in k} ||x_i - m_k||2 \quad (1)$$

Here, we apply K-means clustering to voice files using features collected from them, specifically the average fundamental frequency and average intensity.

IV. RESULTS AND DISCUSSION

We selected speakers who produced 164 speech samples, including short /a/, long /a/, normal /a/, low /a/, and phrase samples. Basic average frequency and strength were selected as the features. Praat, an acoustic analysis program, and scripts written in Praat, which employ the aforementioned standard algorithms to extract pitch and intensity, are used for feature extraction. Graphs are used to display the findings. Python is used to conduct the k-means clustering analysis. Cluster diagrams of the analysed speech data are displayed below. Groups that share a colour cannot share another. The number of clusters used in the analysis ranges from two to four, with k set to those values.

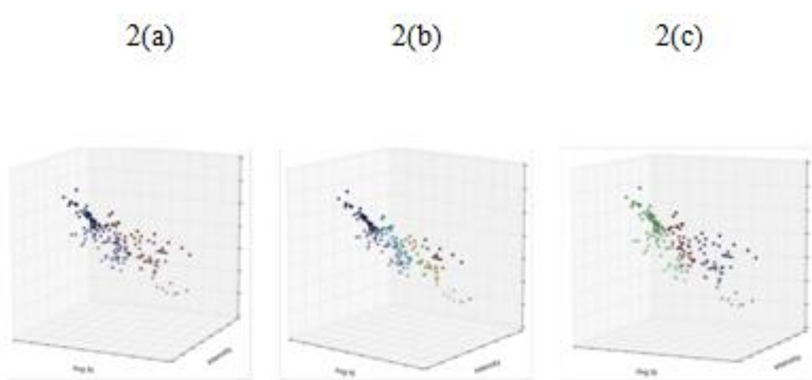

Fig 2 : K-means clustering plot

In Figure 2(a), we have a K-means clustering diagram with two separate clusters, each of which is represented by a different colour. The output is generated by running a Python script on both a Raspberry Pi and an Intel Edison. Figure 2 displays a three-dimensional k-means cluster plot with four distinct clusters indicated by colours (b). The nearest mean is used to determine which cluster an observation belongs to in k-means clustering. We utilised k-means for the feature learning executed in the fog device. In Figure 2 you can see the k-means clustering map for three different clusters, each of which is represented by a different colour (c). The data presented in the figure was produced by running a Python script on a Raspberry Pi and an Intel Edison.

It's possible to use a Raspberry Pi as a cheap computer terminal. The Edison is an embedded computer module designed for use with the Internet of Things. Edison and Raspberry Pi are not equivalent; they have distinct differences in CPU performance and battery consumption. The Machine Learning algorithms were ran on both machines, and the elapsed times, average CPU utilisation, and Memory consumption were compared.

	Edison	**Rpi3**
Run time (ms)	100	5
Max/Min %	10	15
Average %	85	95

Table 1: Intel Edison vs raspberry Pi fog

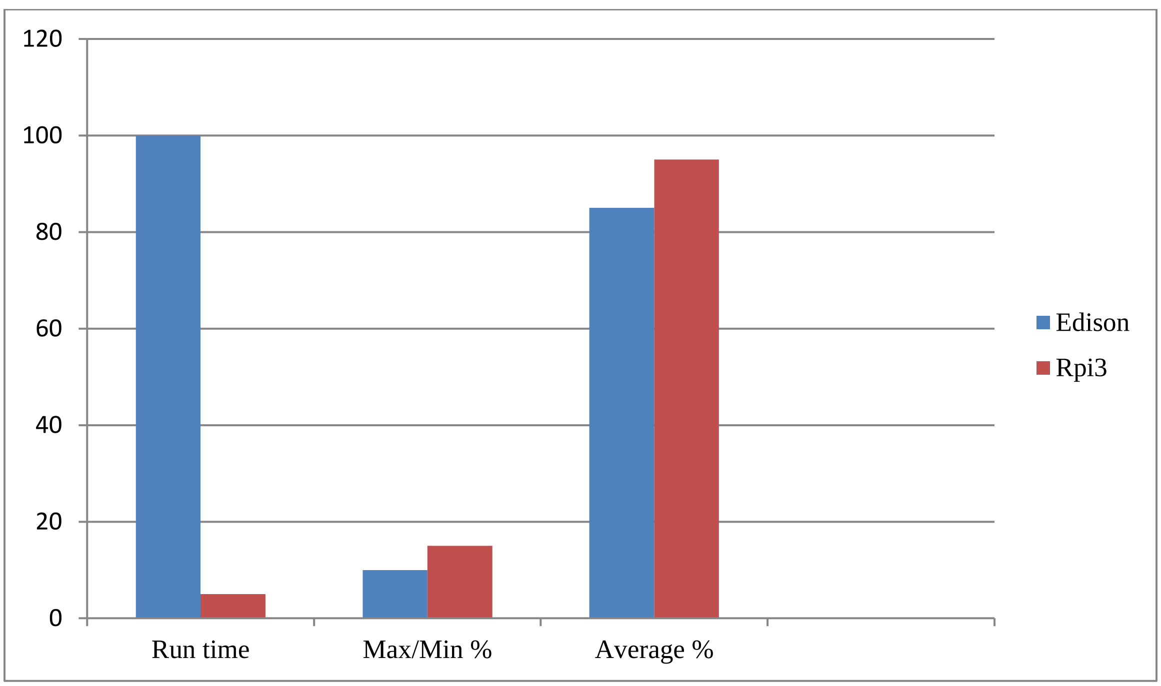

Fig 3: Intel Edison versus raspberry Pi fog relative term

Figure 3 and table 1 depict the parallels and distinctions between the Intel Edison and the Raspberry Pi fog devices. In a perfect system, runtime would be minimised while CPU usage would be maximised and memory usage would be kept to a minimum. The Raspberry Pi either bested the Edison in some areas or performed on par with it. The 200-millisecond real-time response threshold meant that a graphical result could not be generated on the Raspberry Pi for this kind of investigation. Due to the lack of complexity in the required visuals, however, the Raspberry Pi was able to meet the requirement in a mere 160 ms.

V. CONCLUSION

Instead of storing data remotely, as is the case with cloud computing, fog computing keeps data close to where it is really used. This is done alongside local resource pooling, reduced latency, increased QoS, and better user experiences. This article used a Fog computer due to its low resource needs. K-means clustering was used to medical voice data that was gathered from individuals who had Parkinson's disease. The purpose of this was to demonstrate the potential usefulness of K-means clustering (PD). According to the findings of this study, the Smart-Fog framework may prove useful in the assessment and management of medical conditions such as developmental disabilities and therapeutic language understanding within real time. Users of fog computing found that the availability of big data lessened their need to rely on Cloud services. More questions about this proposed design can be asked and answered in the future. It's reasonable to assume that the Fog architecture will play a pivotal role in the near future in shaping how massive data sets are managed and processed.

References

1. Monteiro, H. Dubey, L. Mahler, Q. Yang, and K. Mankodiya, “Fit: A fog computing device for speech tele-treatments,” in IEEE Smart Computing (SMARTCOMP), 2016.
2. S. Hiremath and K. Yang, G.and Mankodiya, “Wearable internet of things: Concept, architectural components and promises for person-centered healthcare,” in Mobihealth Conference. IEEE, 2014.
3. H. Dubey, J. Yang, N. Constant, A. M. Amiri, Q. Yang, and K. Makodiya, “Fog data: Enhancing telehealth big data through fog computing,” in Fifth ASE BigData 2015, Kaohsiung, Taiwan. ACM.
4. J. Andreu-Perez, C. C. Y. Poon, R. D. Merrifield, S. T.C. Wong, and G. Yang, “Big data for health,” IEEE journal of biomedical and health informatics, vol. 19, no. 4, pp. 1193–1208, 2015.
5. R. K. Barik, H. Dubey, C. Misra, D. Borthakur, N. Constant, S. A. Sasane, R. K. Lenka, B. S. P. Mishra, H. Das, and K. Mankodiya, “Fog assisted cloud computing in era of big data and internet-of-things: Systems, architectures and applications,” in Cloud Computing for Optimization: Foundations, Applications, Challenges, p. 23. Springer, 2018.

6. J. Cancela, M. Pastorino, M. T. Arredondo, and O. Hurtado, “A telehealth system for parkinson’s disease remote monitoring. the perform approach,” in 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),, 2013, pp. 7492–7495.