Real-Time Sleep Apnea Diagnosis Method Using Wearable Device without External Sensors

YeongJun Jeon
Department of Electronic Engineering
Kyungpook National University
Daegu, Republic of Korea
thg333@naver.com

KukHo Heo
Department of Electronic Engineering
Kyungpook National University
Daegu, Republic of Korea
hkhpower1@naver.com

Soon Ju Kang
Department of Electronic Engineering
Kyungpook National University
Daegu, Republic of Korea
sjkang@ee.knu.ac.kr

***Abstract*—Currently the diagnosis of sleep apnea is performed mainly in hospital by polysomnography. However, obstructive sleep apnea depend on various factors such as daily life pattern, sleep environment, and posture. Therefore, there is a need for a real-time wearable system that detects sleep apnea which is easy to use. In this paper, we suggest the sleep care system that can predict sleep apnea conveniently whenever wherever. We measured the respiration, SpO2, heartrate, and 3-ACC signals of sleep apnea patients using wearable device. We measured the respiration and SpO2 of patients to judge the levels of sleep apnea. Based on the measurement, we analyzed the heartrate and 3-ACC signals with various machine learning algorithms to determine if sleep apnea correlates with the measurement. As a result of this study, in real-time (640µs), we can diagnosis sleep apnea with 95% accuracy by only analyzing heartrate and 3-ACC signals in a typical smart watch without external sensors.**

Keywords—Machine Learning, Sleep Apnea, KNN, ANN, GNB, Wearable Device, Real-Time, Healthcare

I. Introduction

More than 900 million people worldwide suffer from sleep apnea [1]. The methods for diagnosing sleep apnea are very limited and have several problems. In general, the diagnosis of Sleep Apnea is performed by the polysomnography (PSG) method in the hospital as shown in Figure 1 [2, 3]. In recent years, methods for diagnosing sleep apnea by analyzing PSG results using deep learning have been studied. These studies focus not on real-time diagnosing sleep apnea, but on improving accuracy by processing the results after the measurements are completed [4, 5]. Sleep Apnea symptoms can be difficult to prevent because symptoms are manifested in daily life, such as drinking, smoking, or food intake. In other words, in order to prevent Sleep apnea, it is very important to predict and diagnose the symptoms in real-time in daily life. There was also a study that analyzes the snoring sound signal generated during sleep, but it is judged to be unreliable and inaccurate due to the disturbance of the surrounding environment noise [6]. Furthermore, in order to predict and diagnose sleep apnea in daily life in real-time, a device which is easily to wear is required [7]. However, Personal Sleep Monitoring Device (PMD), a wearable device that measures sleep apnea, is needed to be attach many sensors in daily-life [8, 9]. Therefore, not only it is hard to measure data periodically, but also it is hard for users to analyze the data that has been measured with the device.

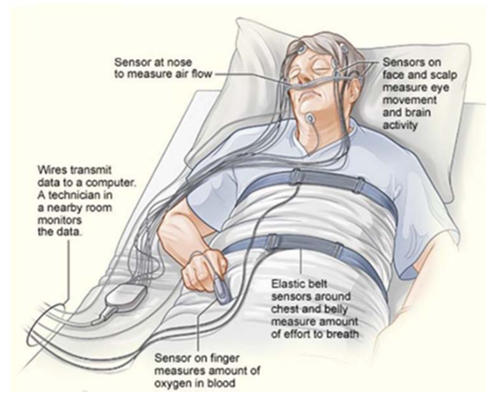

Figure 1 Sleep Polysomnography in hospital: measuring with several sensors attached to the body of the patients.

Therefore, patients with sleep apnea need a wearable system that can diagnosis/predict sleep apnea anywhere and at anytime.

To solve the problem, we suggest the method that can determine sleep apnea in real-time using heartrate and 3-ACC that can be measured mostly with smart watches. First, SpO2, respiration, heartrate and 3-ACC of sleep apnea with patients are measured using sleep care wearable device. Secondly, sleep apnea levels are accurately determined using SpO2 or respiration. Based on the sleep apnea levels, the sleep apnea groups and the normal breathing groups are divided. Finally we use various machine learning algorithms to analyze the correlation between the heart rate and 3-ACC to predict sleep apnea.

The reason why we chose respiration, SpO2, heartrate and 3-ACC is explained: Respiration and SpO2 are major concerns in sleep apnea. The decrease of patient's respiration leads to the decrease of SpO2 which can cause a deadly danger. Heartrate can be easily measured with most of the smart watches out in the market. We use an additional Pulse oximeter with a smart watch to measure heartrate and SpO2 simultaneously. Most of all, it is proven by the research that heartrate can be used to estimate the respiratory value [10, 11]. Moreover, since sleep apnea may occurs depending on postures, we use 3-ACC to sense the wearer's movement. Most sleep apnea patients appear to breathe more frequently when they sleep in supine position than lateral position [12, 13].

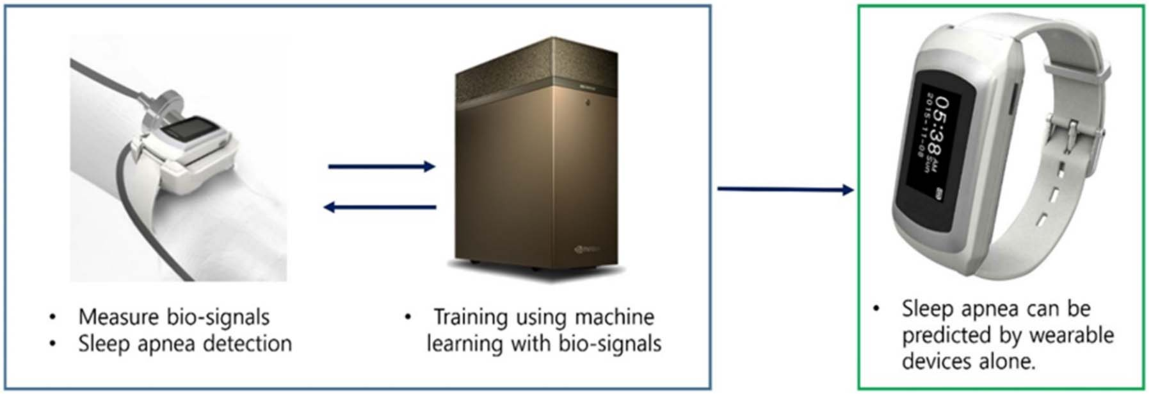

Figure 2 Overview of Suggested System

For real-time analysis of these signals, Gaussian Naive Bayes(GNB), Artificial Neural Network(ANN) and K-Nearest Neighbor(KNN) methods have been tested out. As a result, KNN showed that sleep apnea can be diagnosed with an accuracy of 95% within acceptable timing constraint.

The remainder of this paper is structured as follows. In Section 2, we propose methods and implementation. In Section 3, we introduce the sleep apnea recognition algorithm. In Section 4 the system performance is evaluated. Finally, concluding remarks and plans for our future work are provided in Section 5.

II. Proposed Methods and Implementation

This chapter describes the experimental environment. And we mention the machine learning methods and the criteria for labeling the data.

A. *Overview of Proposed System*

We propose a system that diagnose and predict sleep apnea only with wearable device. To this end, the bio-signals (respiration, SpO2, heartrate and 3-ACC) are measured with a wearable device as shown in Figure 2 [14]. Then we make training data set with 3-ACC, heartrate and label; we designated the label of sleep apnea on the basis of SpO2 and respiration by levels. Finally we analyze the correlation between heart rate and 3-ACC to predict sleep apnea by measuring trained model's accuracy. As a results, it is necessary to wear external sensors for few days to gather learning data, but after that, it can predict sleep apnea using only acceleration and heartrate which can be measured with most of the wearable devices such as smart watch.

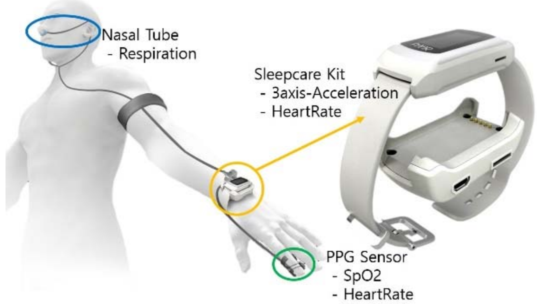

Figure 3 Wearable device configuration & Location of bio-signal measurement

B. *Experience Environment*

In this study, we collect data from sleep apnea patients using a wearable device called Sleep Care Kit(SCK) in Figure 3, which can diagnose sleep apnea in daily life [15] . It can operate for more than 16 hours with a single charge, thus enabling sufficient measurement and real-time analysis during sleep time.

The server environment is NVIDIA DGX Station, Intel Xeon E5-2698 v4 2.2 GHz (20-Core) CPU, and Tesla V100 GPU.

C. *Data Collection*

In this study, we measure five data: respiration, 3-ACC, PPG, heartrate, SpO2 signals; PPG signal is used to calculate the heartrate and SpO2 signals.

We use AASM's sleep apnea judgment criteria as a standard to determine sleep apnea level of a patient with respiration and SpO2 which are measured with SCK[16]. Based on this data, sleep apnea level is determined. According to the result, we analyze whether we can judge sleep apnea level by heartrate and 3-acc signal. The subjects in this study are 5 sleep apnea patients and 3 normal people who are measured for 7 days.

D. *Construction Training Dataset*

We looked at optimal methods for analyzing bio-signals. Analyzing methods are Gaussian Naive Bayes (GNB), Artificial Neural Network (ANN) and K-Nearest Neighbor (KNN).

Machine learning analysis is conducted based on measured data values. Therefore, it is very important to improve the accuracy by selecting the optimal respiration and SpO2 signal levels from the measured data.

Table 1 shows the labeling criterion for respiration signal and SpO2 signal. In Normal group, the respiration values represent the time of continuous breathing without sleep apnea. In contrast, in Apnea group, the respiration values express the time of continuous sleep apnea. The criterion was set by analyzing the point where the data set is divided most evenly and the criteria for judging sleep apnea. Based on AASM's criteria, sleep apnea is prescribed to patients who stops respiration for more than 10 seconds. In this study, the apnea data for apnea group is set as evenly as possible for each labels.

TABLE I. TABLE 1 RESPIRATION & SPO2 LABELING BENCHMARK

	Label	Respiration		SpO2
		Breathing time	Apnea time	
Normal	0	0 ~ 60 sec	Nothingness	100 ~ 97 %
	1	61 ~ 300 sec		96 ~ 95 %
	2	301 sec ~		94 ~ 93 %
Apnea	3	Nothingness	10 ~ 14 sec	92 ~ 90 %
	4		16 ~ 19 sec	89 ~ 86 %
	5		20 ~ 39 sec	85 ~ 0 %

Each of the normal breathing group and apnea group is divided into three groups: 0 to 2 for normal group and 3 to 5 for apnea group. Normal group maintain SpO2 level for more than 96% in daily life. However, the SpO2 level of patients with sleep apnea can fall to 70% or less. In this study, we set normal group as subjects who maintain SpO2 for more than 93% and apnea group as subjects with SpO2 lower than 93%. We specified the levels by dividing each group into 3 sections and labeling them.

III. SLEEP APNEA RECOGNITION ALGORITHM

In this study, the apnea dataset selected to be used in the experiment consists of a data set judged by respiration and a data set judged by SpO2. A respiration data set contains 6,500 data points per label and SpO2 data set contains 9,000 data points per label. These datasets are divided into training datasets and test datasets for 8: 2 ratio, and the training datasets are divided into five cross-validation data sets. The hyper-parameters of each machine learning algorithm are set by using the average of five cross-validation data sets. Finally, we trained the model with the optimal hyper-parameters by using the training data and evaluated the performance of the model by using the test data.

A. Gaussian Naive Bayes (GNB)

GNB is a method that applies the Bayesian algorithm to the 3-acc, SpO2, and heartrate data measured with SCK which is assumed to be a continuous data by the Gaussian distribution. GNB is faster and easier to implement than other machine learning algorithms by constructing Equation (1) and measuring the accuracy using the variance and average of training data [17].

$$P(x_i|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}}\exp\left(-\frac{(x_i-\mu_y)^2}{2\sigma_y^2}\right) \quad (1)$$

In addition, GNB can increase the accuracy by assigning weight to each class, but in this study, we compared the accuracy by using the same weighted model.

B. Artificial Neural Network (ANN)

ANN is consisted of connections between nodes that mimic human neurons. ANN has an input layer, an output layer, and one or more hidden layer. ANN's performance depends on the size of the hidden layer and the number of iterations [18]. Therefore, we applied a Python model in Scikit-learn [19] to construct an input layer using data measured with SCK, and an output layer for labels. In order to improve the performance, we constructed the model by setting the optimal hidden layer (3 layers of equal number of nodes: M) using cross-validation set. We also set the number of iterations to 1000 to minimize the impact on performance and used a learning rate of 0.001, Adam optimizer, and Relu Activation. Then, we tried to find an optimal model. As a result, figure 4 and 5 shows that the accuracy of model is improved when M is increased. The accuracy is converging at a certain point when M was 700 for models using respiration labeled data and when M was 600 for models SpO2 labeled data. Therefore, this paper compares the accuracy with the model using the optimal hidden layer.

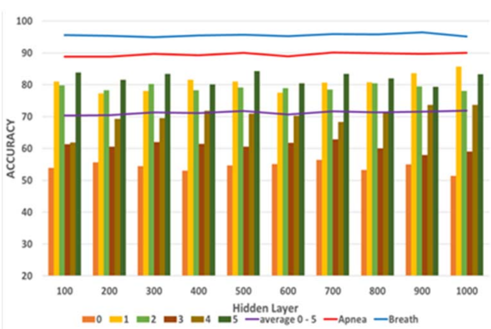

Figure 4 ANN Model using respiration labeled data Accuracy

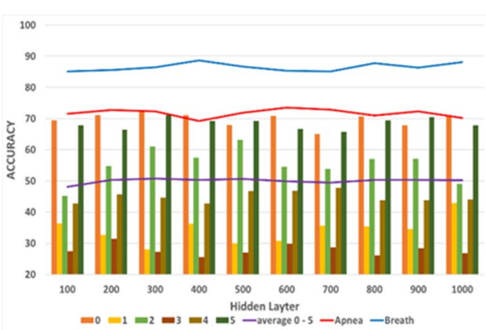

Figure 5 ANN Model using SpO2 labeled data Accuracy

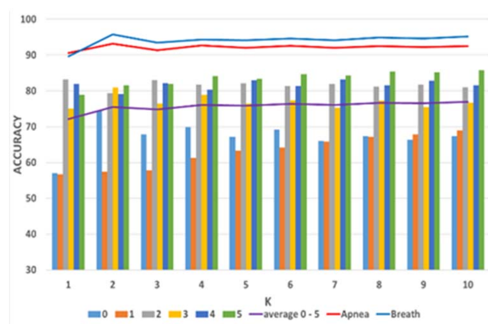

Figure 6 K-NN Model using respiration labeled data Accuracy

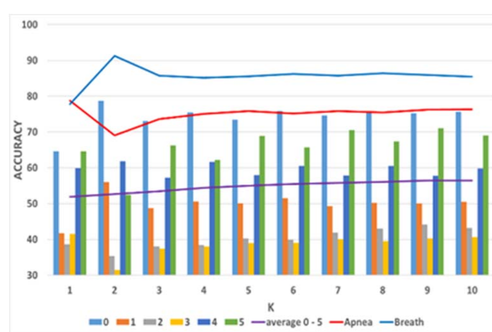

Figure 7 K-NN Model using SpO2 labeled data Accuracy

C. k-Nearest Neighbor (k-NN)

The k-NN is finding k training data with the highest similarity to the measured data, and then classifies it as the group that occupies the largest weight. The advantage of the k-NN classifier is that it is possible to improve accuracy by adjusting the number of k. Therefore, for the high performance model, the data measured in the SCK were divided into 5 cross-validation sets and then experimented to find the optimal k value. As a result, figure 6 and 7 shows that when k is 1, respiration / apnea performance of the model using respiration labeled data, according to Err (1-NN) $\leqslant 2 \times$ IdealErr (error of the ideal model that can fit the given data) [20] certification, is about 90%, and respiration / apnea performance of the model using SpO2 reference data is about 78%. In addition, we found out the best performance by testing out different values of k from 1 to 10 for each model: k=2 for respiration labeled data and k=4 for SpO2 labeled data. Therefore, this paper compared the accuracy with the best model.

IV. Performance Evaluation

Since we aim to process in real-time, we measured execution time to demonstrate the effectiveness of the study. Furthermore, we discuss the best way to diagnosis sleep apnea and its accuracy.

A. Measurement Time Evaluation

In the future, we aim for a system that can analyze the data measured every 40 ㎳ on a wearable device to predict sleep apnea risk in real-time. Therefore, to do this, the data must be learned and the results should be derived in less than 40 ㎳ in real-time. As shown in Table 2, the execution speed of each algorithm in this paper is as follows.

The GNB model takes 242.4 ㎲ at respiration labeled data and 244.1 ㎲ at SpO2 labeled data. With the use of the variance and the mean of training data from Equation (1), GNB is the fastest model among tested algorithms.

The ANN model takes 274.5 ㎲ at respiration labeled data and 264.1 ㎲ at SpO2 labeled data. The performance speed of ANN depends on the number of features and how the hidden layer is constructed. The hidden layer is composed of 3 layers with 600 nodes each in this ANN models. Therefore, speed is guaranteed because the model is not huge.

The KNN model takes 640.7 ㎲ at respiration labeled data and 638.7 ㎲ at SpO2 labeled data. Because KNN compares measured data with all learning data, performance time increases in accordance with the data size. Therefore, as data increases, the performance time gets slower compared to other algorithms [21]. Thus, it has been supplemented by many recent studies and guarantees speed in low computing environment [22].

Based on these results, the system we designed, which focuses both on accurate results and short measuring time within 40 ㎳, is more suitable for real-time Apnea detection than other existing systems.

B. Accuracy Evaluation

In this study, we measured the data and analyzed the patients' daily living environment. The analysis assessed the accuracy of the weekly measured data which excludes abnormal conditions (sensor dropout, toilet movement during measurement). Table 3 compares the accuracy of each machine learning algorithms. SpO2 is less accurate in determining sleep apnea because the response occurs 10 to 20 seconds after the apnea situation occurs. Therefore, in this paper, data analysis by respiration sensor value is considered as a meaningful result.

1) GNB

The GNB has an accuracy of about 89% to distinguish respiration / apnea. The reason for this is that in the case of a data structure that is rotated like an accelerometer, even though the first value and the last value are almost equal, there appears to be a large error, which results in increased variance and lower accuracy [17].

In addition, 3-acc and heartrate data have small in variance of each class, but the difference between the means of each class is small, resulting in lower accuracy.

2) ANN

The ANN shows an accuracy of about 71% for each class and 92% for normal respiration and apnea. In the case of ANN, better performance can be obtained by increasing the depth of the hidden layer [9].

3) KNN

It can be seen that KNN is able to analyze 74% of each class, 94% of apnea and normal respiration. The KNN algorithm compares similarity measured data to all learning data, and classifies it as the group that occupies the largest weight. Therefore, the accuracy is highest. According to KNN's characteristic, Currently, 95% accuracy is measured with weekly data, but if more data is carried out, the accuracy is expected to be higher.

TABLE II. TABLE 2 Average Performance Time of Machine Learning

Label	GNB	ANN	KNN
Breath	242.4 ㎲	274.5 ㎲	640.7 ㎲
SpO2	244.1 ㎲	264.1 ㎲	638.7 ㎲

TABLE III. TABLE 3 Performance of machine learning algorithms

	Breathing (%)			SpO2 (%)		
Model	GNB	ANN	KNN	GNB	ANN	KNN
Label 0	21.5	50.4	72.7	20.3	73.1	75.9
Label 1	97.3	87.1	55.4	4.0	40.7	53.2
Label 2	67.5	81.8	80.8	90.3	62.2	39.5
Label 3	7.8	59.2	80.7	2.7	26.9	36.3
Label 4	88.9	64.8	78.2	11.4	49.8	61.7
Label 5	5.3	84.7	81.1	56.1	71.8	62.4
Average 0 - 5	48.0	71.3	74.8	30.8	54.1	54.8
Apnea	79.7	87.6	92.3	50.2	73.0	73.4
Breath	99.	97.3	95.8	76.4	89.1	87.2

Therefore, we will carry out research to cure sleep apnea by using KNN algorithm that can assure high accuracy in real-time in embedded environment.

V. Conclusion & Further Work

This study discusses a method for diagnosing sleep apnea conveniently in daily life. To solve the problem, we proposed diagnosis method of sleep apnea using only the heartrate and 3-ACC signals that can be measured on smartwatch without external sensors. Furthermore, we apply K-NN algorithm to a measured data from each user and personalize it to predict sleep apnea in real-time. In this paper, it is shown that sleep apnea with K-NN algorithm can be predicted only with heartrate and 3-acc signals in 640 ㎲ time with about 95% accuracy. This will enable real-time diagnosis of sleep apnea in most smart watches.

Now, we are developing the smart pillow. So we are not just about diagnosing sleep apnea, we are doing research that can solve sleep apnea.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2018R1A6A1A03025109)

References

[1] "175 Million Europeans Have Sleep Apnea, Highlighting Scale of Global Health Crisis." [Online]. Available: https://newsroom.resmed.com/175-million-europeans-sleep-apnea

[2] A. Westerlund, Y. T. Lagerros, G. Kecklund, J. Axelsson, and T. Åkerstedt, "Relationships between questionnaire ratings of sleep quality and polysomnography in healthy adults," *Behavioral sleep medicine,* vol. 14, no. 2, pp. 185-199, 2016.

[3] "File:Sleep studies.jpg." https://commons.wikimedia.org/wiki/File:Sleep_studies.jpg (accessed.

[4] D. Dey, S. Chaudhuri, and S. Munshi, "Obstructive sleep apnoea detection using convolutional neural network based deep learning framework," *Biomedical engineering letters,* vol. 8, no. 1, pp. 95-100, 2018.

[5] O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih, and U. R. Acharya, "Deep learning for healthcare applications based on physiological signals: A review," *Computer methods and programs in biomedicine,* vol. 161, pp. 1-13, 2018.

[6] A. S. Karunajeewa, U. R. Abeyratne, and C. Hukins, "Multi-feature snore sound analysis in obstructive sleep apnea–hypopnea syndrome," *Physiological measurement,* vol. 32, no. 1, p. 83, 2010.

[7] L. Haoyu, L. Jianxing, N. Arunkumar, A. F. Hussein, and M. M. Jaber, "An IoMT cloud-based real time sleep apnea detection scheme by using the SpO2 estimation supported by heart rate variability," *Future Generation Computer Systems,* vol. 98, pp. 69-77, 2019.

[8] H. J. Kim, "Portable sleep monitoring devices in Korea," *Korean Journal of Otorhinolaryngology-Head and Neck Surgery,* vol. 56, no. 2, pp. 68-73, 2013.

[9] R. Santos-Silva *et al.*, "Validation of a portable monitoring system for the diagnosis of obstructive sleep apnea syndrome," *Sleep,* vol. 32, no. 5, pp. 629-636, 2009.

[10] A. Garde, W. Karlen, J. M. Ansermino, and G. A. Dumont, "Estimating respiratory and heart rates from the correntropy spectral density of the photoplethysmogram," *PloS one,* vol. 9, no. 1, p. e86427, 2014.

[11] W. Karlen, S. Raman, J. M. Ansermino, and G. A. Dumont, "Multiparameter respiratory rate estimation from the photoplethysmogram," *IEEE Transactions on Biomedical Engineering,* vol. 60, no. 7, pp. 1946-1953, 2013.

[12] L. Sonnesen, "Head posture and upper cervical spine morphology in patients with obstructive sleep apnea," in *Sleep Apnea-Recent Updates*: InTech, 2017.

[13] R. D. McEvoy, D. J. Sharp, and A. T. Thornton, "The effects of posture on obstructive sleep apnea," *American review of respiratory disease,* vol. 133, no. 4, pp. 662-666, 1986.

[14] "NVIDIA DGX STATION." nvidia.com. https://www.nvidia.com/en-us/data-center/dgx-station/ (accessed.

[15] Y. J. Jeon and S. J. Kang, "Wearable Sleepcare Kit: Analysis and Prevention of Sleep Apnea Symptoms in Real-Time," *IEEE Access,* vol. 7, pp. 60634-60649, 2019.

[16] "2.Obstructive Sleep Apnea (OSA): Practice Essentials, Background, Pathophysiology. ." Emedicine.medscape.com. https://emedicine.medscape.com/article/295807-overview#a2. (accessed.

[17] H. Zhang, "The optimality of naive Bayes," *AA,* vol. 1, no. 2, p. 3, 2004.

[18] J. Gracia, A. Mazon, and I. Zamora, "Best ANN structures for fault location in single-and double-circuit transmission lines," *IEEE Transactions on Power Delivery,* vol. 20, no. 4, pp. 2389-2395, 2005.

[19] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *Journal of machine learning research,* vol. 12, no. Oct, pp. 2825-2830, 2011.

[20] A. Kontorovich and R. Weiss, "A Bayes consistent 1-NN classifier," in *Artificial Intelligence and Statistics*, 2015, pp. 480-488.

[21] K. H. Heo, S. Y. Jeong, and S. J. Kang, "Real-time user identification and behavior prediction based on foot-pad recognition," *Sensors,* vol. 19, no. 13, p. 2899, 2019.

[22] Z. Chen and J. Yan, "Fast KNN search for big data with set compression tree and best bin first," in *2016 2nd International Conference on Cloud Computing and Internet of Things (CCIOT)*, 2016: IEEE, pp. 97-100.