Received November 21, 2021, accepted December 11, 2021, date of publication December 24, 2021,
date of current version January 12, 2022.

Digital Object Identifier 10.1109/ACCESS.2021.3138460

A Novel Data-Driven Adaptive Technique to Generate a Physician Visiting Schedule for Better Patient Experience

SARIA SAFDAR, SHOAB AHMED KHAN, AND ARSLAN SHAUKAT
Department of Computer and Software Engineering, College of Electrical and Mechanical Engineering (CEME), National University of Sciences and Technology (NUST), Islamabad 44000, Pakistan

Corresponding author: Saria Safdar (sariasafdar@yahoo.com)

ABSTRACT This paper proposes an adaptive data-driven novel technique for generating physician's schedules exploring the trade-off between physician waiting time and patient experience. Generally, in hospitals fix time slots are assigned to the patients, without differentiating patients into different categories. Some patients are referral patients who need less time, others are first-time visitors who need more time than referral patients. Some are serious patients who take even more time, so a fixed time slot can increase the waiting time both at the doctor and the patient level. So, there is a need to categorize patients according to their treatment which could help in improving the waiting time of doctors and as well of the patients. This technique trains the patient's data according to their specific category based on the time doctors usually spend on them. This can help in generating a schedule that will minimize the waiting of patients and at the same time improve the hospital performance by scheduling maximum patients ensuring that the same patient's get the best experience. To achieve this a data-driven scheduling algorithm (DDSA) is proposed which uses probability distribution, clustering and doctor's scheduling algorithm to classify patients into different categories. The data is collected through the RFID machines. The patients are given cluster head time for the treatment. Different patient categories are given different treatment times. Then patient and doctor commutative waiting times are calculated. The system applies different set-up times to allow hospital management to trade-off between doctor and patient waiting times. The average waiting time for each patient using DDSA comes out to be seven minutes. Hence, this technique can help hospitals across the globe in improving their performance and patient experience.

INDEX TERMS Clustering, data driven, probability distribution, patient experience, RFID, waiting times.

I. INTRODUCTION

The healthcare services of every country is a sensitive and prioritized concern in a person's life. Any deviation from providing primary healthcare services is a major concern. Currently, most of the hospitals are overcrowded and lack efficient patient management systems [1], which result in long queues and increased patient waiting times [2]. These waiting times plays an important role in calculating the patient experience [3]. Patient experience is a multi-dimensional paradigm covering a number of features of care. The most common areas covering the patient experience are appointment scheduling, waiting times, long queues, attitude and courtesy of staff, provision of lab reports, cleanliness, information assistance by nurses and treatment by doctors [4]. The feedback on these areas identifies the weakness and strengths which is clearly reflected by the patient rating. These ratings mostly focus on what happened in these areas, which service needs improvement, which service is meeting the satisfactory threshold etc. [5], [6]. Patient feedback ratings are routinely used to measure the quality of hospital services [7]. There are multiple ways to collect these experiences, most commonly adopted techniques are surveys, interviews, email, telephone calls etc. Good experience is directly related to the profitability of the hospital, and bad patient experience increases the retention rates [8]. Many studies have shown an increased mortality rate due to long queues at various stations of the hospital. The normal approach followed in the hospitals is to give the same time limit/slot to all the patients, for this patients are called in bulk, which increases

The associate editor coordinating the review of this manuscript and approving it for publication was Xi Peng.

TABLE 1. Waiting time in Various Hospitals [9].

Hospital	Range(mins)	Mean(mins)
Hospital 1	28-70	51
Hospital 2	24-64	45
Hospital 3	27-62	46
Hospital 4	25-70	49
Hospital 5	26-66	50
Hospital 6	50-95	70
Hospital 7	35-86	62
Hospital 8	33-83	59

the frustration levels as the patients have to wait a lot for their turn.

A. PROBLEM DOMAIN

The treatment time consumption of every patient, at each station, does not lie in the same time slot range. It depends upon the condition and category of the patient. For example, in the case of an MRI scan, the time slot required for an old age patient is generally longer than the young patient. Fig. 1 depicts the normal scenario followed in the hospitals. All the patients are given same times 20 minutes for the treatment which is shown as the blue line, but in reality every patient takes a different treatment time, shown as the orange line. Therefore, some patients may take more than 20 minutes and some patients may take less than 20 minutes. Hence in this paper, we identified and calculated different treatment times for different patients based on their category. Table 1 shows the study conducted regarding the waiting times at eight different hospitals [9]. It is clear from the table that a patient has to wait a lot for his treatment.

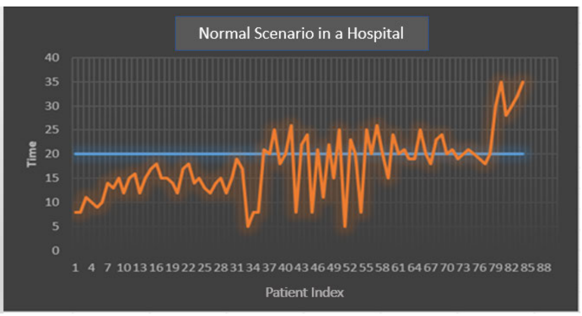

FIGURE 1. Normal scenario of treatment time given to patients in hospitals.

Various cases have been filed stating that the mortality rate increased due to waiting times at different stations of the hospital. Fig.2 shows a case study of women who died because of waiting time associated with the registration counter. The criticality of the patient was ignored and being asked to get the admission slip first before getting the treatment. The women died waiting in the queue at the registration counter due to overcrowding and manual work performed at the registration station. The other reason for being overcrowded is the limited number of staff working at these counters which


PAKISTAN TODAY

15 MUST READ ARTICLES | HEADLINES | NATIONAL | CITY | BUSINESS | FOREIGN | ENTERTAINMENT | SPORTS | COMMENT

DG Khan woman dies while waiting for hospital admission slip

BY MONITORING REPORT , (LAST UPDATED OCTOBER 10, 2018)

Following hours of wait in queue to get a hospital admission slip for treatment, a young woman died at a Teaching Hospital in Dera Ghazi Khan. According to *reports*, the doctors allegedly refused to treat her over the failure to present the admission slip.

A resident of Mozah Jhariwala, a village near Drahma, the woman, wife of Abdul Aziz was brought to the emergency ward of the teaching hospital with a complaint of heart pain. However, instead of administering immediate medical checkup, the doctors on duty sent the patient to the OPD department where she was forced to wait in the long queue for two hours to get the admission slip.

As her condition deteriorated, she collapsed on the ground and died, following which her family protested against the hospital administration and blamed them for the

FIGURE 2. Case study of a women who dies waiting in long queue.


Thousands of patients die waiting for beds in hospitals - study

Doctors' report finds 5,449 deaths since 2016 followed delayed admission to A&E

Latest election news - live updates

FIGURE 3. Increase in mortality rate due to long waiting times [10].


THE CONVERSATION

Academic rigour, journalistic flair

Arts + Culture | Business + Economy | Education | Environment + Energy | Health + Medicine | Politics + Society

Wait times remain stubbornly long in hospital emergency rooms

April 29, 2020 10.12pm AEST

FIGURE 4. Long waiting times [11].

cannot handle unexpected patients, resulting in long queues and increased waiting times.

In [10]–[12], different case studies were presented regarding the mortality rate associated with long waiting times as shown in Fig.3 and 4.

Therefore, there is a need for an automated system that can reduce these waiting times both at the doctor level and at the patient level. To achieve this a data-driven scheduling algorithm (DDSA) is developed based on the probability distributions, clustering and doctor's scheduling algorithm. The DDSA divides the patients into four categories namely report showing, follow-up, new patient and serious patient, using a clustering algorithm. The identified patient is given the cluster head time for the treatment. Then the doctor and the patient waiting times are calculated. After that, different set-up times are applied to reduce the doctors waiting time. The system allows the management to set the threshold based on the number of patients and doctors. The system also provides a trade-off between doctor and patient waiting times.

B. CONTRIBUTIONS

1) The proposed framework applies the probability distributions to the timestamp data of various stations to check the high time variability and then applies clustering algorithm that divides the patients into four categories; report showing, follow up, new patients and serious patients.Normally patients are not categorized and are given the same treatment time.
2) The framework identifies and calculates different treatment times for different patients based on the identified category using a Data-Driven Scheduling Algorithm (DDSA).
3) The proposed framework develops a doctor schedule to reduce doctor and patient waiting times by applying different set-up times. It allows the management to set the best trade-off between doctor and patient waiting time according to the overcrowding at diffrent stations.

The rest of the paper is organized as follows: section II describes the literature survey, section III describes the proposed framework in detail, section IV summarized the results and section V concludes the overall research work.

II. LITERATURE SURVEY

Chen *et al.*, predicted the waiting time of each patient's treatment task by developing a Patient Treatment Time Prediction (PTTP) algorithm based on a random forest algorithm. Based on this approach hospital recommendation system for queue management was developed. The motivation for the developed system was the overcrowding in the hospital and lack of queue management [13]. Vollmer *et al.*, compared different machine learning algorithms with time series analysis for the prediction of overcrowding in the emergency department of the hospital to keep the process running in a smooth way. This prediction helped the hospital management to kept a right number of staff at the emergency department. The stacked regression is used to compare the results of both methods [14]. Roopa *et al.*, used a random forest algorithm based PTTP model to predict the treatment time of each task. After calculating the treatment time required by each task a hospital(QRS) was developed to efficiently manage the time for each patient thus reducing the waiting time [15].

Kumar *et al.*, discussed the issues of long queues, overcrowding and long waiting times, that resulted in patient dissatisfaction for the particular hospital. Keeping these issues in mind, the authors developed a time prediction algorithm to efficiently manage different treatment tasks time. The patients with different dependent tasks were identified. The developed system worked as follows: treatment time of each task is also calculated using the average time spent at each task. The algorithm then calculated the overall time required for each patient and also the waiting time required for each treatment task in the hospital. A mobile-based application using android technology is developed. GSM is used for notification in android phones. The connection is secured using AES encryption algorithm [16].

Permanasari *et al.*, developed a decision support system for the inpatient room availability. The developed system predicted the recovery time of each admitted patient so that new patients based on the predicted time of discharge can be admitted maintaining stability in the hospital. The ID3 algorithm was used to build the decision tree which is further classified into classification rules. The estimated recovery time was divided into three categories between 1-10 days. The classification accuracy of the system was 60% [17].

Handayani *et al.*, developed a system for outpatient treatment time prediction. The data of 120 patients was collected from the eye clinic of the Cirebon hospital. The time prediction based on a random forest algorithm, first classifies data based on gender and age, and then predicted a waiting time based on these attributes. The results were shown using gender/age regression tree. According to the results of the developed system, it is observed that age plays an important role in treatment time. Early ages have less treatment time than the old ages due to morbidity issues [18].

Javadifard *et al.*, discussed the issues of long waiting times in the phlebotomy units, as all the patients from other outpatient departments visits this unit for their tests. So, there is a need for a proper tests time plan for each patient. To achieve the proper plan ANN was used. The sigmoid function and cross-entropy was used. The accuracy of the system was 88% [19].

Patil *et al.*, used a random forest algorithm for the prediction of the waiting time of patients at different departments of the hospital. The patient's find it convenient, if they know the waiting time associated with each task. Patients can go for another test, if one test's waiting time is longer than the other. The authors developed a recommendation system with the minimized path awareness.The system determined and forecasted a well-organized and suitable management plan suggested for every patient [20].

Savanth *et al.*, proposed an efficient queuing system for hospitals based on the treatment tasks time. The issue of long waiting time frustrates the patients. To solve this issue an advanced treatment time and waiting time plan for each task is determined and conveyed to the patients. The steps of the developed algorithm are: calculate the waiting time for each patient, then sort the tasks in ascending order but if the tasks are dependent then time is calculated based on their dependencies instead of waiting time. In the end, based on the calculated time treatment plan is given to the patient [21].

Jin *et al.* in [22], investigated the issues of prolonged waiting time associated with an eye clinic in Singapore. It is observed that delay in arrival time is one of the main reasons for the long waits. FlexSim healthcare was used as a discrete simulation model that simulates the time for each patient in real-time. The input to the model was the difference in arrival pattern, the difference in scheduling new appointments in the already set appointments for the day and the different process flow. The output of the simulation model was the waiting time associate with each patient.

Luo *et al.*, predicted the number of patients that arrives to the emergency department in the China hospital, using time series analysis. The results helped the hospital management to be well aware of the risk of overcrowding and long waits [23]. In [24], Liu proposed a smart phone- based prediction system using a case-based reasoning algorithm (CBR). The application helped the patients to reach the hospital at the specified time, therefore reducing the wait time.

Hamza *et al.*, developed a simulation-based model for patient flow (SIM-PEED) in the hospital. The techniques used to develop the system included ''discrete event simulation, agent-based simulation,a multi-attribute decision-making method'' [25]. The output of the system showed less waiting time and shorter patient stay in the hospital.

Osman *et al.* in [26], developed a Bluetooth low energy-based real-time location awareness system integrated with the hospital information system. The visualization of patients helped the hospital management to reduce the overcrowding and the long waiting times.

In [27], the authors emphasized organizing the patient flow in the emergency department. Quality Function Deployment (QFD), was used to reduce the waiting time and overcrowding in the hospitals by organizing the patients according to their treatments.

Weerakoon *et al.*, in [28], highlighted the issues of patients frustration because of long waits and overcrowding at each station (registration, vitals and doctor). The authors proposed a queuing theory analysis using two parameters arrival time and service time. The theory calculated the mean waiting time for a week and based on similar cases assigned treatment time to new patients. The system worked on a first come first serve basis.

A. CLUSTERING USED IN HEALTHCARE SYSTEMS

Rani *et al.*, developed a diabetes prediction system using association clustering and time series. The steps of the developed system are: the training set of data is collected, it is serialized based on the chronological order, then clustering is applied and association rules are developed which helped in prediction. The system is helpful as it warned the patients beforehand of the future probability of the disease based on historical attributes. Time series analysis is compared with ANN. The results showed system has better accuracy with time series analysis than ANN [29].

TABLE 2. The year-wise detail of the techniques used to reduce patient waiting times. The table also mentioned the dataset, the technique applied and the limitation.

Author	Year	Techniques applied	Dataset used	Limitations
Hamza et al	2021	SIM-Peed Simulation based model on similar cases	Patient waiting time	Time given based on similar cases. Patients not categorized
Osman et al	2021	Bluetooth low energy based real time location awareness with HIS	Patients dataset	Data visualization based overcrowding reduction
Vollmer et al	2021	Time Series Analysis	Patients dataset	Emergency department patients catered
Handayani et al	2020	Random Forest Algorithm	Patients dataset	Time given based on age
Savanth et al	2017	Queuing System based on Time series	Patients Dataset	Average waiting time for each task is calculated beforehand
Chen et al	2016	Patient Treatment Time Prediction (PTTP), Random Forest Algorithm	Patients Dataset	Hospital recommendation system for queue management was developed
Adel et al	2018	Quality Function Deployment (QFD)	Patients dataset	Patient categorized only in emergency department
Patil et al	2019	Hospital Queuing Recommendation system based on minimum path awareness using Random Forest Algorithm	Patients dataset	Treatment time for each task is calculated beforehand and plan is conveyed to the patient
Weerakon et al	2019	Queuing theory FCFS basis	Patients Dataset	Time given based on similar case analysis
Javadifard et al	2019	ANN is used. Sigmoid function is used as an activation function and cross entropy as a loss function.	Patients Dataset	Average time is calculated for each task
Liu et al	2014	Case-based reasoning (CBR) algorithm	Patients Dataset	Time given based on similar cases in past Patients not categorized
S. Safdar et al	2021	Data Driven based scheduling algorithm using Clustering , distributions and doctors scheduling algorithm	Patients timestamp Dataset	The system categorizes the patients into four groups namely report showing, follow up,new patients and serious patients. The time is given according to the category of the patients. This reduces the waiting time of both the patients and the doctor

Marne *et al.*, developed a breast cancer prediction system that helped in the early detection of disease and treatment. Wisconsin breast cancer dataset was used. The data was clustered using K means clustering that grouped closely related attributes. After the data is divided into groups, the decision tree is applied to make association rules. These rules predicted the presence or absence of disease [30].

In [31], Kumar *et al.*, predicted the presence of tumour using clustering techniques. X-means and K-mean performance is measured. The results showed K-means performed well than X-means. Umamaheswari *et al.*, developed a prediction system for myocardial infarction. The data was collected from UCI with 270 records. The data wass preprocessed using normalization. Then relief feature selection method applied. Initially, there were 18 features but after feature selection, reduced to the 8 most frequent features. Clustering was performed using the K-means algorithm. The clustering results were further used to predict heart disease [32].

B. RESEARCH GAPS

From the literature it is observed that in hospitals, time is given to patients based on similar cases of previous days, patients are not categorized regarding the treatment and most of the hospitals are using data visualization based overcrowding reduction. Some hospitals are giving treatment time based on age. Furthermore, the average waiting time for each task is calculated beforehand and then treatment time is given. Mostly, the same treatment time is given to all the categories of patients. To the best of our knowledge, in previous studies patients are not categorized while giving the treatment time. The patients have to wait for their turn at different stations. There is a need for an automated system that uses sensors to record patient time at different stations to reduce manual work and then categorizes patients for treatment time.

III. PROPOSED FRAMEWORK

The flow chart in Fig. 5 shows the patients journey in the hospital. There are multiple steps that a patient has to follow before a doctor can see him/her. At each station, the patient has to either wait for the service or is standing in the queues for the registration. This waiting time bothers the patients a lot especially in the case of emergency, where the patient needs quick treatment. The flowchart in Fig. 6 shows the wait and queue time associated with each station. There is a need for an automated system that reduces this wait time at each station and allows the hospital service management to adjust the waiting times according to the situation of the patients and the overcrowding. The patients should be intimated about the wait times beforehand. To achieve this, a data-driven scheduling algorithm (DDSA) is developed. This algorithm develops a doctors schedule in which the doctor and patients waiting times are reduced to the minimum.

Fig. 7 shows the proposed framework. The data for DDSA is collected through RFIDs deployed at three stations (registration, nursing station and doctor) of the hospital. Each patient is given an RFID card that he swipes on the station RFID machine. Three timing information of each patient is saved at each station [33]. This timing information is used by DDSA for further analysis.

A. DATA-DRIVEN SCHEDULING ALGORITHM (DDSA)

The steps involved in the development of the DDSA are:

1) Distributions
2) Clustering
3) Doctor Schedule

B. DISTRIBUTIONS

The first step in the development of the DDSA is the probability distributions. We have applied different distributions to the timestamp data of the patients to check the time variability at each station. This variability is represented through mean and variance.

We have applied different distributions on the following stations:

1) Registration Station
2) Vital Station
3) Doctor Station

1) REGISTRATION STATION

When different probability distributions are applied on the registration station data as shown in Fig. 8. The Poisson distribution fitted well as shown in Fig. 9. The X-axis represented by data in the figure shows the time given to different patients and the Y-axis represented as density shows the probability distribution function (pdf).

Where,
RNP -Registration Normal/already registered Patients
RNewP - Registration New Patients
AVG -Average

Fig. 9 shows that the patients normally takes 3 minutes at the registration station. The patients below average are represented through "RNP" are normal patients and the patients which are above the average are new patients represented through "RNewP".

2) VITALS STATIONS

When different distributions are applied to the vitals station Gaussian distribution fitted well to the data. Fig. 10 shows the result of Gaussian distribution on the vitals time data. X-axis shows the time given to different patients and Y-axis shows the probability distribution function.

Where,
NV-Normal Vitals
NVNP- Normal Vitals new patients
AVG-Average time

Fig. 11 shows that the average time the patients take at the vitals station is 5 minutes. The patients which are below average are mostly those patients who have normal vitals checking procedures and the patients who are above the average are mostly those patients who are new or serious patients whose vitals checking process takes more time than the usual patients.

3) TIME DISTRIBUTION ON DOCTOR STATION

The treatment time taken by different patients at doctor station is between 4-36 minutes. We have applied different probability distributions on the patient treatment time data at doctors station. Fig. 12 shows the result of Normal distribution on the doctor time, as it is seen normal distribution is not covering the data completely so we applied more probability distributions.

Fig. 13 shows the result of Gamma distribution on the doctor's time and it can be seen that the distribution fitted well.Fig. 14 shows the comparison of the Gamma and Gaussian distribution on the doctor's time. It can be clearly seen in the Fig. 14 that Gamma distribution fitted well.

Where,
RS-Report Showing
F-UP- Follow Up
Avg- Average

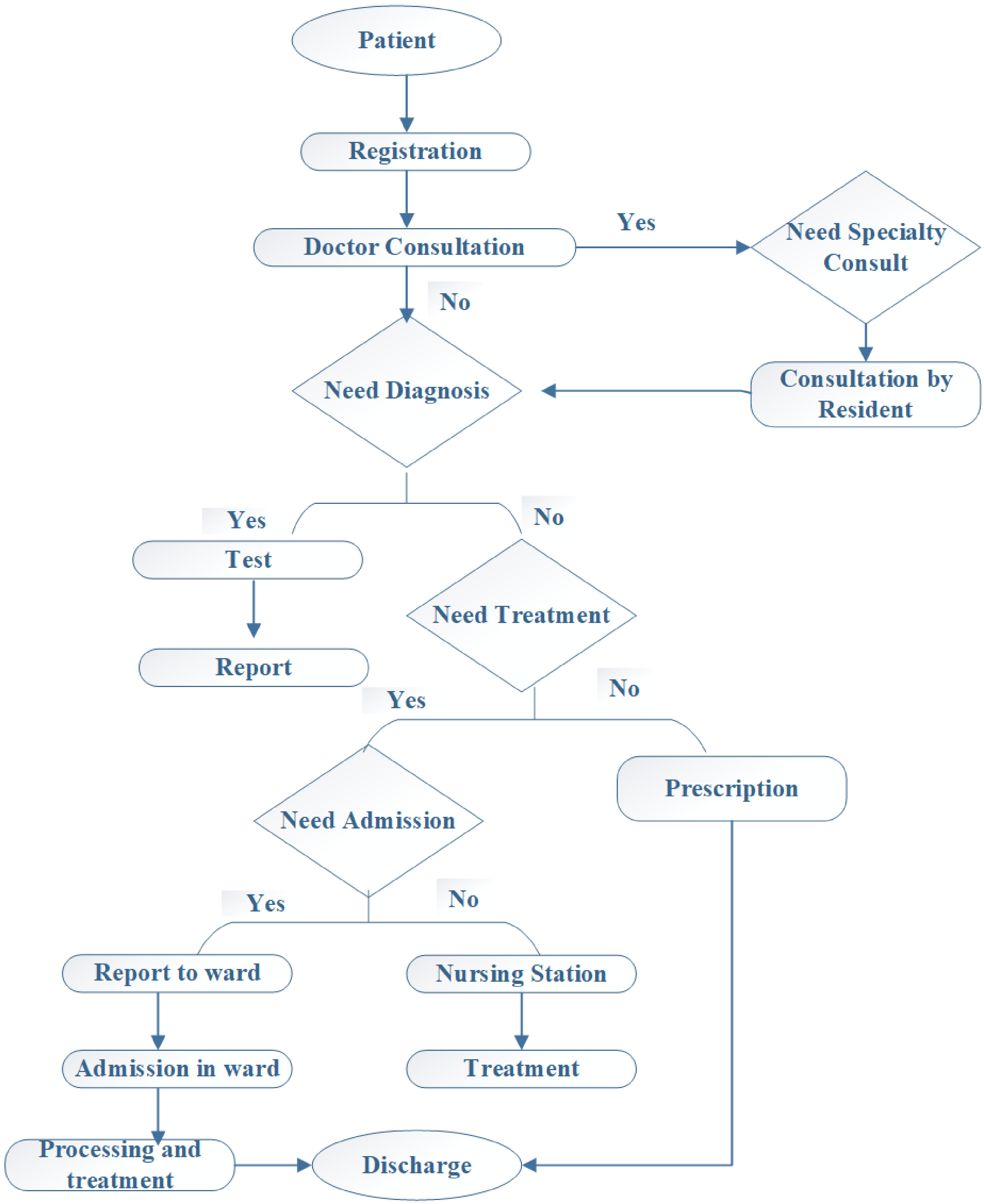

FIGURE 5. Patient journey in the hospital.

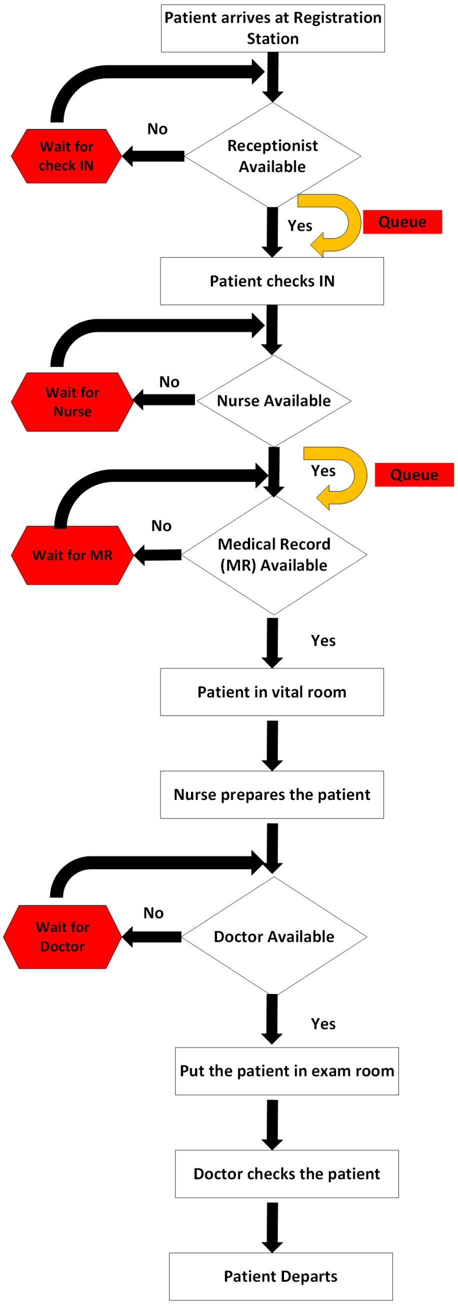

FIGURE 6. Wait and queue time associated with patients journey in the hospital.

NP- New Patient

SP- Serious Patient

Fig. 15 shows that every patient is taking different treatment times and the variance at the doctor station is high which compels us to further investigate that which types of patients are taking less time and which are taking more time. The patients which are below average are mostly those patients who come up for report showing or follow up and the patients who are above the average are mostly those patients who are new or serious patients.

C. CLUSTERING

The K-means clustering algorithm is applied to the timestamp data of the patients at the doctor station to form the natural clusters. The high variance at the doctor's station shows that there are different groups of patients who take different times. Fig. 16 shows the clustering framework.

Fig.17 shows the spread of the patient's treatment time is between 4 to 36 minutes.The X-axis shows the time taken by patients and the Y-axis shows the patient index. Fig. 18 shows the result of the k-means clustering algorithm with the four centroids. The algorithm divided the timestamp data into four clusters with cluster 1 ranging from 4 to 8 minutes, cluster 2 ranging from 11 to 15 minutes, cluster 3 ranges from 16 to 22 minutes and cluster 4 ranging from 25 to 35 minutes as shown in Fig. 19.

After the natural clusters are made new patient data is used for prediction. When new data is given to the k-means algorithm, each data points distance from all the centroids is calculated using the Euclidean distance formula.The data point whose distance is minimum with the centroid among the four centroids, that data point is placed in that respective centroid's cluster. Likewise, all the data points distances are calculated and are placed in their respective clusters as shown in Fig. 20.

After the natural clusters are made we analyzed the common attributes of each cluster. Based on the common attributes the clusters are given names by consulting the doctor as shown in Fig. 16 (clustering framework). Cluster 1 which takes the less time of all the clusters is named as,

"Report Showing" as the most common attributes are the blood and urine reports. The time taken by cluster 1 patients is from 4-8 minutes. Cluster 2 is named as,

"Follow Up" based on the most common attributes which are age and co morbidity (multiple issues). The time taken by cluster 2 patients is from 9-13 minutes. The 3rd cluster is named as,

"New Patients" and the common attributes are hypertension, high cholesterol, stress and diabetes. The time taken by cluster 3 patients is 14 to 21 minutes. The 4th cluster is named as,

"Serious Patients" with common attributes as stroke, trauma, high blood sugar and MI. The time taken by cluster 4 patients is 22-36 minutes.

D. DOCTORS SCHEDULE

The doctor's schedule is developed using the "scheduling algorithm" which has the following steps.

1) DOCTOR'S SCHEDULE ALGORITHM (DSA)

1) *Identify patient type*
   (report showing, follow-up, new patient, serious patient)
2) *Assign cluster to the identified patient type*
   C1 = report showing
   C2=follow up

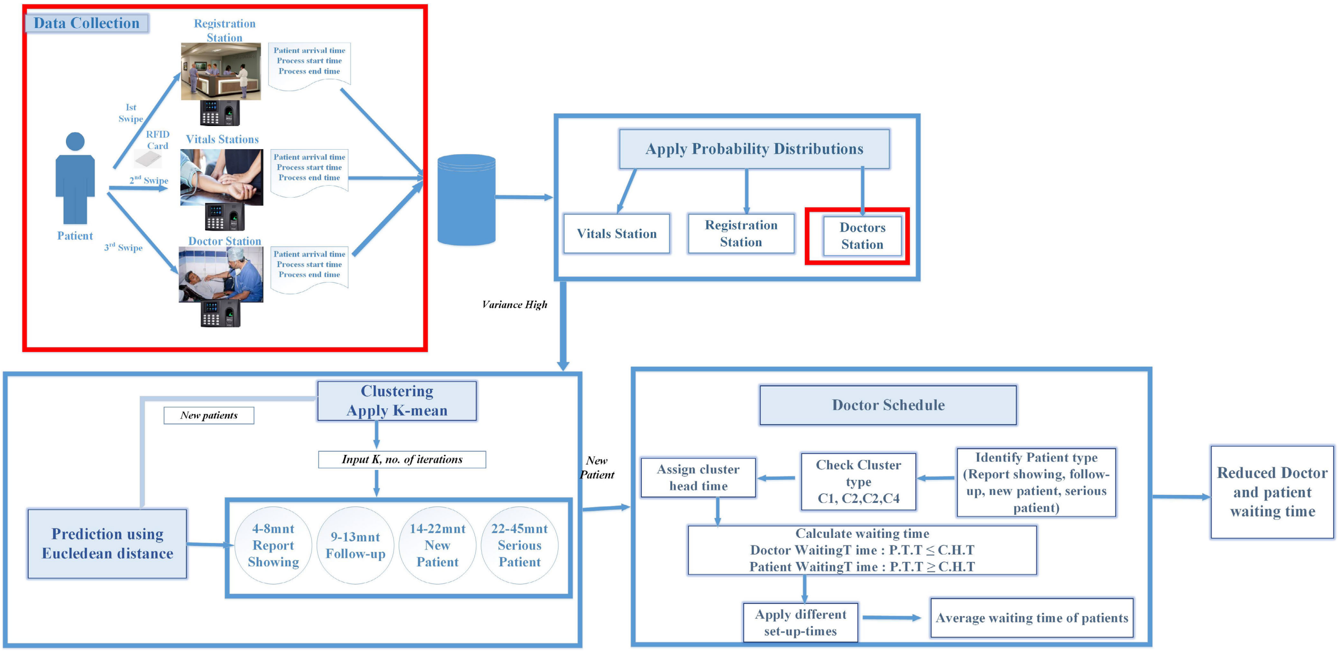

FIGURE 7. Framework of Data Driven Scheduling Algorithm (DDSA).

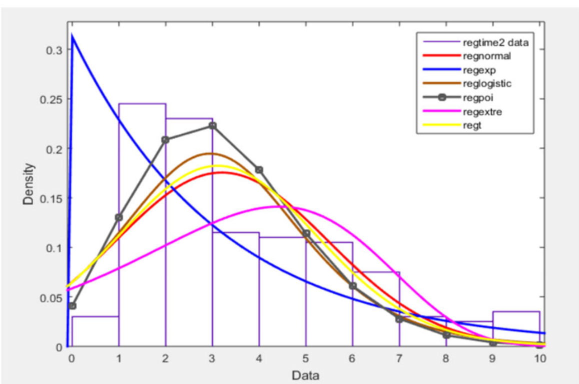

FIGURE 8. Result of different distributions applied to the registration data.

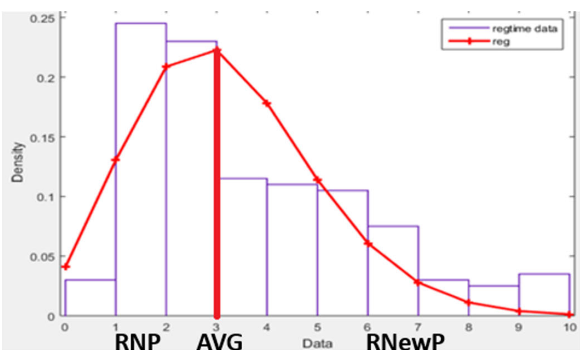

FIGURE 9. Patient categorization based on time at registration station.

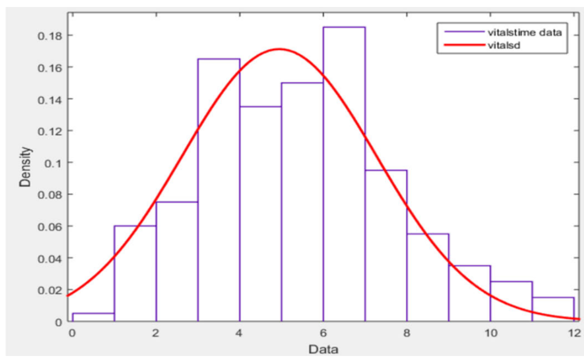

FIGURE 10. Normal distribution fitted well on Vitals station timestamp data.

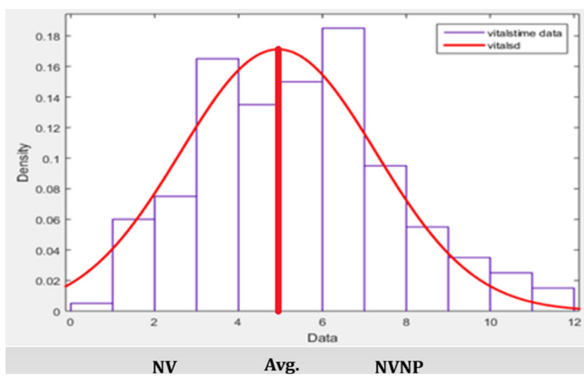

FIGURE 11. Patient categorization based on time at vitals station.

C3=new patient
C4=serious patient
Where, C1,C2,C3,C4 are the clusters

3) *Assign cluster head time according to cluster type*
4) *Calculate waiting times (doctor waiting time and patient waiting time) using the formulas*

$$DoctorWaitingTime : P.T.T \leq C.H.T \tag{1}$$

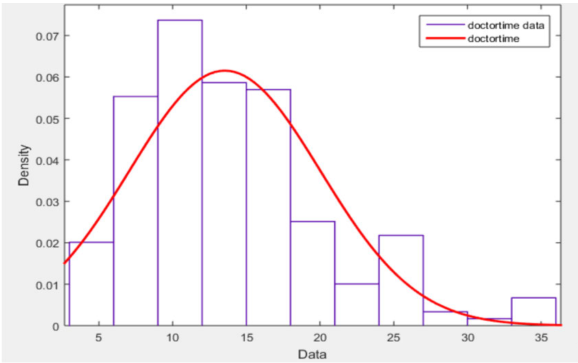

FIGURE 12. Normal distribution on doctors timestamp data.

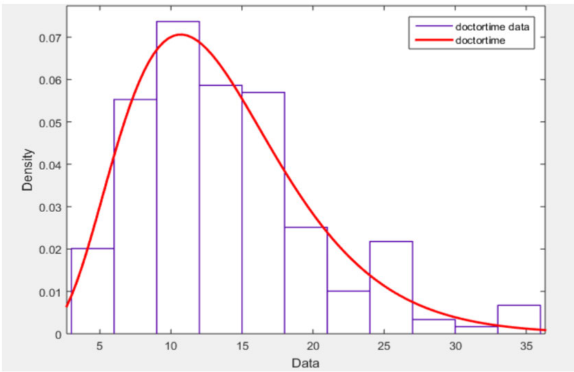

FIGURE 13. Gamma distribution on doctors timestamp data.

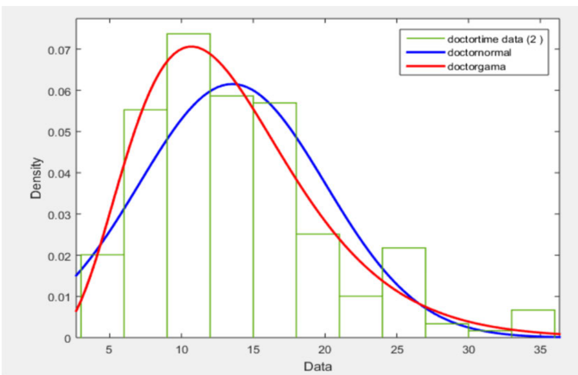

FIGURE 14. Comparison of Gamma distribution and normal distribution on doctors station.

$$PatientWaitingTime : P.T.T \geq C.H.T \tag{2}$$

5) *Reduce waiting times by setting different setup times*
6) *Compute average patient waiting time*

$$AverageP.W.T = \sum_{i=1}^{n} \frac{P.W.T}{N} \tag{3}$$

where,
P.T.T is patient treatment time
C.H.T is cluster head time

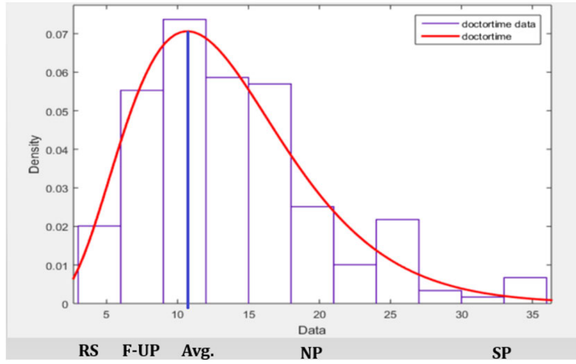

FIGURE 15. Patient categorization based on time using gamma distribution at doctors station.

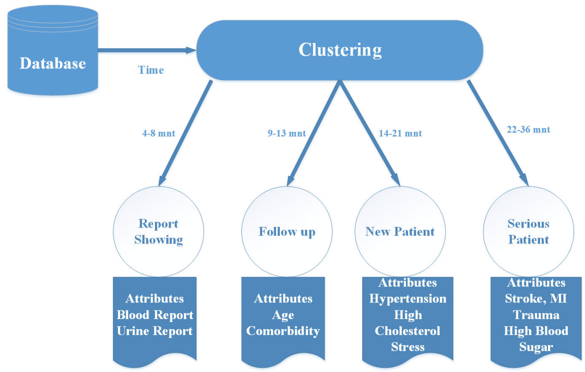

FIGURE 16. Clustering framework: The figure shows the four clusters with time and common attributes.

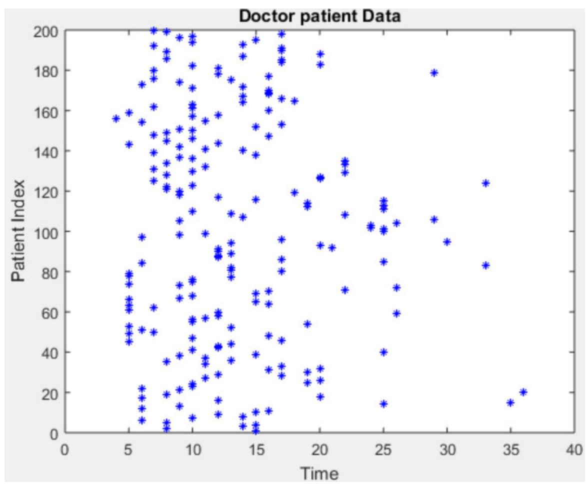

FIGURE 17. Spread of patient time data.

P.W.T is patient waiting time
N is the number of patients

IV. RESULTS

A. CASE STUDY 1

When a patient arrives at the registration station the system first checks about the type of the appointment i.e. if a

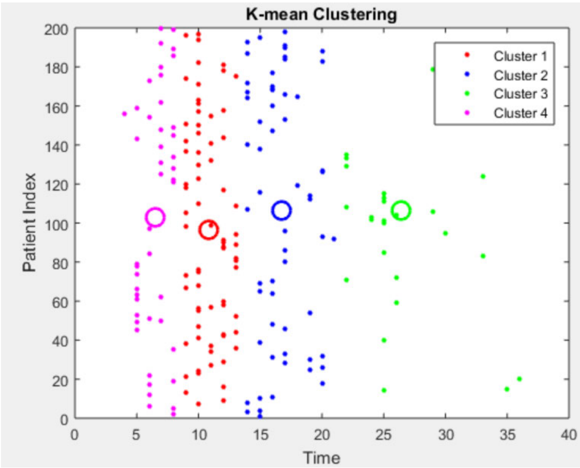

FIGURE 18. K-Means Clustering Results.

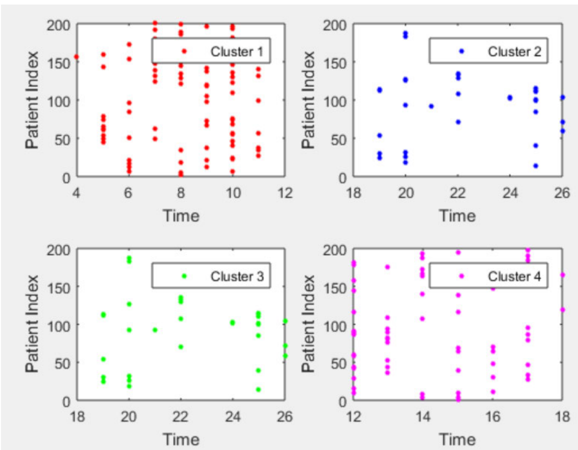

FIGURE 19. Sub graphs of Clusters using K-Means: Cluster 1 shows the report showing patients, cluster 2 shows the follow-up patients, cluster 3 shows the new patients and cluster 4 shows the serious patients.

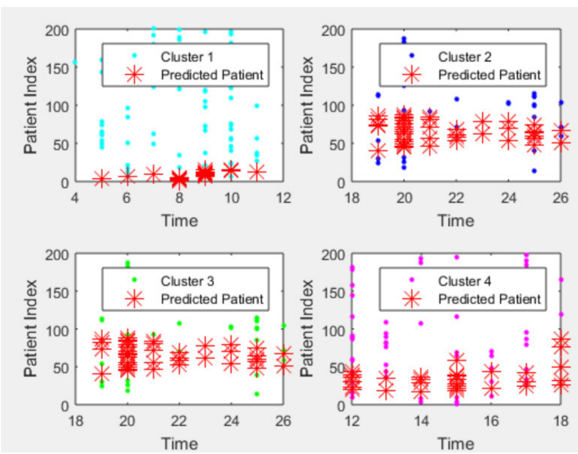

FIGURE 20. Prediction results.

patient comes for the report showing, follow up, new patient or serious patient. Once the type of the patient is known the algorithm puts the patient in his respective cluster and gives him the cluster centroid time. For example, if a patient comes for a report showing then the algorithm puts the patient in C1 (cluster 1) and gives him 7 minutes as the treatment time. Likewise, all the different types of patients are placed in their respective clusters. Afterwards patient and doctor waiting times are calculated and different set-up times are set to reduce these waiting times. For case study 1, 25 patients are considered. Table 3 shows the details of case study 1. A.T.T is the actual time taken by the patients, D.W is the doctor waiting and P.W is the patient waiting time.

Once all the appointments are given centroid times w.r.t their clusters, then the patient visits the doctor clinic according to their turns. For example, patient 1 who come up for a report showing is placed in cluster 1 and is given cluster head time of 7 minutes for the treatment. The patient goes into the doctor's clinic and comes out taking only 6 minutes of the specified time. The remaining 1 minute is put into the doctor waiting time calculated using equation 1. If a patient takes more than the specified cluster head time, that extra time is put into the patient waiting time calculated using equation 2. So we acquired two types of waiting time from the data, doctors waiting time and patients waiting time.

$$DoctorWaitingTime : P.T.T \leq C.H.T \tag{4}$$
$$PatientWaitingTime : P.T.T \geq C.H.T \tag{5}$$

where,

P.T.T is Patient Treatment Time

C.H.T is Cluster Head Time/ Cluster Centroid Time

After calculating the waiting times of all the individual patients a cumulative sum is calculated at the end as shown in table 4. The table shows some values for the doctor waiting time and patient waiting time. The algorithm then set-up different set-up times to reduce these waiting times.

1) DIFFERENT SET-UP TIMES

Set-Up Time 3: The DSA starts with 3 minutes as the initial set-up time. The appointment schedule starts 3 minutes earlier than the given appointment time. This will reduce the doctors waiting time, but the patient's waiting time increases as the patients are called 3 minutes earlier than the scheduled appointment. Table 5 shows the result of doctor and patient waiting times after the set-up time 3. It can be clearly seen in table 6 that the doctor's waiting time is reduced but patient waiting time is increased.

Set-Up Time 4: The set-up time is increased to 4 minutes. The patients are called 4 minutes earlier than the given appointment schedule time.This set-up time more reduces the doctor's waiting time but increases the patients waiting time as they are called 4 minutes earlier than the appointment time. Tables 7 and 8 shows the results of set-up time 4. As it can be clearly seen doctors waiting time is reduced but patients

TABLE 3. Case study 1: The number of patient in this case study are 25.

id	Patients Type	Cluster	Cluster head	A.T.T	D.W	P.W
1	Report Showing	C1	6	4	2	0
2	Report Showing	C1	6	12	0	6
3	Follow up	C2	11	11	0	0
4	Follow up	C2	11	14	0	3
5	Report Showing	C1	6	5	1	0
6	Follow up	C2	11	14	0	3
7	Follow up	C2	11	10	1	0
8	Serious patient	C4	27	23	4	0
9	Serious patient	C4	27	41	0	14
10	Report Showing	C1	6	5	1	0
11	Report Showing	C1	6	8	0	2
12	Follow up	C2	11	10	1	0
13	Follow up	C2	11	13	0	2
14	Follow up	C2	11	8	3	0
15	Serious patient	C4	27	20	7	0
16	New Patient	C3	17	25	0	8
17	New Patient	C3	17	22	0	5
18	Serious patient	C4	27	24	3	0
19	Follow up	C2	11	15	0	3
20	New Patient	C3	17	26	0	9
21	New Patient	C3	17	10	7	0
22	New Patient	C3	17	15	2	0
23	New Patient	C3	17	22	0	5
24	Serious patient	C4	27	15	12	0
25	Serious patient	C4	27	24	4	0

TABLE 4. Cumulative doctor and patient waiting time.

Total Doctor Waiting Time	Total Patient Waiting Time
47	61

TABLE 5. Doctor and patient waiting times using set-up time 3.

Id	D.W	P.W
1	0	0
2	0	9
3	0	3
4	0	6
5	0	3
6	0	6
7	0	3
8	1	3
9	0	17
10	0	3
11	0	5
12	0	3
13	0	5
14	0	3
15	4	3

waiting time is increased. So to reduce this doctors waiting time we applied set-up time 5.

Set-up time 5: Now the set-up time is set to 5 minutes. Hence, we can see that by applying different set-up times the doctor waiting time has become zero and the average patient time is seven minutes. Tables 9 and 10 shows the results of set-up time 5 minutes. When the doctors waiting time comes out to be zero we calculated the average waiting time for each patient using equation 6.

$$AverageP.W.T = \sum_{i=1}^{n} \frac{P.W.T}{N} = \frac{181}{25} = 7minutes \qquad (6)$$

TABLE 6. Cumulative doctor and patient waiting times using set-up time 3.

Total Doctor Waiting Time	Total Patient Waiting Time
18	133

TABLE 7. Doctor and patient waiting times using set-up time 4.

Id	D.W	P.W
1	0	0
2	0	10
3	0	4
4	0	7
5	0	4
6	0	7
7	0	4
8	0	4
9	0	18
10	0	4
11	0	6
12	0	4
13	0	6
14	0	4
15	0	4

TABLE 8. Cumulative Doctor and patient waiting times using set-up time 4.

Total Doctor Waiting Time	Total Patient Waiting Time
5	157

TABLE 9. Doctor and patient waiting times using set-up time 5.

Id	D.W	P.W
1	0	0
2	0	11
3	0	5
4	0	8
5	0	5
6	0	8
7	0	5
8	0	5
9	0	19
10	0	5
11	0	7
12	0	5
13	0	7
14	0	5
15	0	5

TABLE 10. Cumulative Doctor and patient waiting times using set-up time 5.

Total Doctor Waiting Time	Total Patient Waiting Time
0	181

TABLE 11. Summary of different waiting times using case study 1.

Waiting Times	Initial W.T	S.U.T 3	S.U.T 4	S.U.T 5
Doctor Waiting Time	47	18	5	0
Patient Waiting Time	61	133	157	181

Table 11 shows the summary of different waiting times applied to case study 1.

Where,

Initial W.T is the initial waiting time

S.U.T 3 is the set-up time 3

TABLE 12. Case study 2:The number of patients in this case study are 25.

Id	Patient Type	Cluster	Cluster head	A.T.T	D.W	P.W
1	Follow up	c2	11	9	2	0
2	serious patient	c4	27	23	4	0
3	Follow up	c2	11	11	0	0
4	report showing	c1	6	5	1	0
5	New Patient	c3	17	25	0	8
6	Follow up	c2	11	12	0	1
7	New Patient	c3	17	20	0	3
8	Follow up	c2	11	10	1	0
9	serious patient	c4	27	45	0	18
10	Follow up	c2	11	12	0	1
11	New Patient	c3	17	24	0	7
12	serious patient	c4	27	27	0	0
13	Follow up	c2	11	13	0	2
14	New Patient	c3	17	10	7	0
15	Follow up	c2	11	9	2	0
16	serious patient	c4	27	37	0	10
17	Follow up	c2	11	17	0	6
18	report showing	c1	6	8	0	2
19	report showing	c1	6	10	0	4
20	Follow up	c2	11	4	7	0
21	New Patient	c3	17	23	0	6
22	report showing	c1	6	6	0	0
23	serious patient	c4	27	6	21	0
24	Follow up	c2	11	8	3	0
25	New Patient	c3	17	16	1	0

TABLE 13. Summary of different waiting times using case study 2.

Waiting Times	Initial W.T	S.U.T 4	S.U.T 5	S.U.T 6
Doctor Waiting Time	49	23	12	6
Patient Waiting Time	68	164	188	212

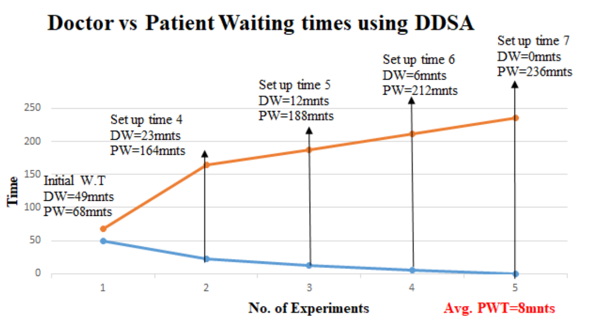

FIGURE 21. Trade-off graph between doctor and patient waiting times.

S.U.T 4 is the set-up time 4
S.U.T 5 is the set-up time 5

B. CASE STUDY 2

For case study 2, again 25 patients are selected with different treatment times(actual time taken) as shown in table 12. Different set-up times are applied to check where doctor's waiting time comes out to be zero. For this purpose we performed three experiments with set-up times 4, 5 and 6. Table 13 shows the summary of different set-up times applied to the case study 2 data. At the end average patient waiting time is calculated using equation 7, which in this case is around 8 minutes.

$$AverageP.W.T = \sum_{i=1}^{n} \frac{P.W.T}{N} = \frac{212}{25} = 8minutes \qquad (7)$$

The graph in Fig. 21 shows the best trade-off between doctor waiting time and patient waiting time. The orange line shows the patient waiting time and the blue line shows the doctor waiting time. Different set-up times are applied to achieve the minimum waiting time for the patients. The management can choose the trade-off between the waiting times according to the situation.

V. CONCLUSION

Healthcare services quality is a global and prioritized concern [34], [35]. Patient experience is an evaluation metric that is used to check the quality of a hospital [36]. This experience is not only related to the medical treatment, rather multiple factors are involved in its evaluation. Among many factors waiting time is the prioritized metric. Overcrowding's in hospitals with poor patient management systems is the main cause of long waiting times and long queues. All the patients are given the same treatment times and are called in bulk for the services. The patient through this approach has to wait for his turn which increases the frustration levels. Therefore, there is a need for an automated system that reduces these waiting times both at the doctor and patient levels. To achieve this, a novel framework is proposed that introduces an adaptive data-driven scheduling algorithm. The framework is developed using probability distributions, clustering and a doctor's scheduling algorithm. The patients are divided into four categories using clustering. Then each patient is given the centroid time as the treatment time. Through this approach, all the patients are given different treatment times. The system also allows the management to set different thresholds (patient and doctor waiting times) according to the doctor and patient availability. The average patient waiting time at the doctor's station using the proposed framework is below 10 minutes, while the average waiting time as mentioned in table 1 is above 40 minutes. The proposed framework using DDSA will help the hospital management with timely response to real-time issues regarding waiting time and long queues. This will help to maintain a long term hospital-patient loyalty and decreases the patients rate of leaving the particular hospital due to low patient experience ratings. The proposed framework deals with out-patients timestamp data only, the same framework can be applied to in-patients and the emergency patients by defining the new categories using clustering and then treatment time is given to those categories by using DDSA.

REFERENCES

[1] M. E. K. Ajoud and T. A. Jouili, "Healthcare service quality in government hospitals: Evaluating patients' satisfaction," *Calitatea*, vol. 22, no. 182, pp. 106–109, 2021.

[2] C. Bleustein, D. B. Rothschild, A. Valen, E. Valatis, L. Schweitzer, and R. Jones, "Wait times, patient satisfaction scores, and the perception of care," *Amer. J. Managed Care*, vol. 20, no. 5, pp. 393–400, 2014.

[3] E. Larson, J. Sharma, M. A. Bohren, and O. Tunçalp, "When the patient is the expert: Measuring patient experience and satisfaction with care," *Bull. World Health Org.*, vol. 97, no. 8, p. 563, 2019.

[4] S. A. Cosma, M. Bota, C. Fleáeriu, C. Morgovan, M. Văleanu, and D. Cosma, "Measuring Patients' perception and satisfaction with the Romanian healthcare system," *Sustainability*, vol. 12, no. 4, p. 1612, Feb. 2020.

[5] R. Z. Q. Almomani, R. R. Al-Ghdabi, and K. M. Banyhamdan, "Patients' satisfaction of health service quality in public hospitals: A PubHosQual analysis," *Manage. Sci. Lett.*, pp. 1803–1812, 2020.

[6] M. Umoke, P. C. I. Umoke, I. O. Nwimo, C. A. Nwalieji, R. N. Onwe, N. Emmanuel Ifeanyi, and A. Samson Olaoluwa, "Patients' satisfaction with quality of care in general hospitals in Ebonyi State, Nigeria, using SERVQUAL theory," *SAGE Open Med.*, vol. 8, Jul. 2020, Art. no. 2050312120945129.

[7] B. Berkowitz, "The patient experience and patient satisfaction: Measurement of a complex dynamic," *Online J. Issues Nursing*, vol. 21, no. 1, Jan. 2016, doi: 10.3912/OJIN.Vol21No01Man01.

[8] L. H. Aiken, D. M. Sloane, J. Ball, L. Bruyneel, A. M. Rafferty, and P. Griffiths, "Patient satisfaction with hospital care and nurses in england: An observational study," *BMJ Open*, vol. 8, no. 1, Jan. 2018, Art. no. e019189.

[9] M. Z. Haroon and I. H. Thaver, "Patient flow and waiting time in emergency department of tertiary health care hospitals of khyber pakhtunkhwa," *J. Mech. Continua Math. Sci.*, vol. 14, no. 6, pp. 484–497, Nov./Dec. 2019, doi: 10.26782/jmcms.2019.12.00033.

[10] D. A. Martinez, H. Zhang, M. Bastias, F. Feijoo, J. Hinson, R. Martinez, J. Dunstan, S. Levin, and D. Prieto, "Prolonged wait time is associated with increased mortality for Chilean waiting list patients with non-prioritized conditions," *BMC Public Health*, vol. 19, no. 1, pp. 1–11, 2019.

[11] P. K. Plunkett, D. G. Byrne, T. Breslin, K. Bennett, and B. Silke, "Increasing wait times predict increasing mortality for emergency medical admissions," *Eur. J. Emergency Med.*, vol. 18, no. 4, pp. 192–196, Aug. 2011.

[12] K. A. Melon, D. White, and J. Rankin, "Beat the clock! Wait times and the production of 'quality' in emergency departments," *Nursing Philosophy*, vol. 14, no. 3, pp. 223–237, Jul. 2013.

[13] J. Chen, K. Li, Z. Tang, K. Bilal, and K. Li, "A parallel patient treatment time prediction algorithm and its applications in hospital queuing-recommendation in a big data environment," *IEEE Access*, vol. 4, pp. 1767–1783, 2016.

[14] M. A. C. Vollmer, B. Glampson, T. Mellan, S. Mishra, L. Mercuri, C. Costello, R. Klaber, G. Cooke, S. Flaxman, and S. Bhatt, "A unified machine learning approach to time series forecasting applied to demand at emergency departments," *BMC Emergency Med.*, vol. 21, no. 1, pp. 1–14, Dec. 2021.

[15] G. Roopa and C. Nirmala, "A hospital queuing-recommender system for predicting patient treatment time using random forest algorithm," *Int. J. Eng. Manuf. Sci.*, vol. 8, no. 1, pp. 119–133, 2018.

[16] M. A. Kumar, S. Mahesh, T. Kamalraj, and M. Azhagiri, "Patient treatment time prediction for out patient department OPD," *Int. J. Eng. Sci. Comput.*, vol. 7, no. 4, pp. 6238–6242, 2017.

[17] A. E. Permanasari, A. D. W. Ikram, M. N. Aji, H. A. Nugroho, and I. S. Sakkinah, "A web-based decision support system of patient time prediction using iterative dichotomiser 3 algorithm," in *Proc. 11th Int. Conf. Inf. Technol. Electr. Eng. (ICITEE)*, Oct. 2019, pp. 1–6.

[18] D. Handayani, M. Mustafid, and B. Surarso, "Patient queue systems in hospital using patient treatment time prediction algorithm," *Kinetik, Game Technol., Inf. Syst., Comput. Netw., Comput., Electron., Control*, vol. 5, no. 1, 2020, doi: 10.22219/kinetik.v5i1.1001.

[19] H. Javadifard, S. Sevinc, O. Yildirim, D. Orbatu, E. Yasar, and A. R. Sisman, "Predicting patient waiting time in phlebotomy units using a deep learning method," in *Proc. Innov. Intell. Syst. Appl. Conf. (ASYU)*, Oct. 2019, pp. 1–4.

[20] P. Patil and S. Thakur, "Patient waiting time prediction in hospital queuing system using improved random forest in big data," in *Proc. Int. Conf. Issues Challenges Intell. Comput. Techn. (ICICT)*, Sep. 2019, pp. 1–10.

[21] S. S. Savanth and K. N. R. Mohan Babu, "Hospital queuing-recommendation system based on patient treatment time," in *Proc. Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, Jun. 2017, pp. 953–958.

[22] X. Jin, A. I. Sivakumar, and S. Y. Lim, "A simulation based analysis on reducing patient waiting time for consultation in an outpatient eye clinic," in *Proc. Winter Simulations Conf. (WSC)*, 2013, pp. 2192–2203.

[23] L. Luo and Y. Feng, "Using time series analysis to forecast emergency patient arrivals in CT department," in *Proc. 12th Int. Conf. Service Syst. Service Manage. (ICSSSM)*, Jun. 2015, pp. 1–5.

[24] Y. T. Liu, Y. H. Chen, N. P. Wang, C. J. Lin, and Y. A. Chen, "Reducing the outpatient waiting time based on CBR algorithm using smartphone," in *Proc. Int. Conf. Mach. Learn. Cybern. (ICMLC)*, vol. 2, Jul. 2014, pp. 506–511.

[25] N. Hamza, M. A. Majid, and F. Hujainah, "Sim-pfed: A simulation-based decision making model of patient flow for improving patient throughput time in emergency department," *IEEE Access*, vol. 9, pp. 103419–103439, 2021.

[26] M. S. Osman, A. Azizan, K. N. Hassan, H. A. Ghani, N. H. Hassan, F. Yakub, S. M. Daud, and L. A. Latiff, "BLE-based real-time location system integration with hospital information system to reduce patient waiting time," in *Proc. Int. Conf. Electr., Commun., Comput. Eng. (ICECCE)*, Jun. 2021, pp. 1–6.

[27] H. Adel, M. A. Wahed, and N. Saleh, "A novel approach for improving patient flow in emergency department," in *Proc. 9th Cairo Int. Biomed. Eng. Conf. (CIBEC)*, Dec. 2018, pp. 69–72.

[28] W. M. N. B. Weerakoon, S. Vasanthapriyan, and U. A. P. Ishanka, "A queuing model for outpatient department to reduce unnecessary waiting times," in *Proc. 14th Conf. Ind. Inf. Syst. (ICIIS)*, Dec. 2019, pp. 203–208.

[29] S. Rani and S. Kautish, "Association clustering and time series based data mining in continuous data for diabetes prediction," in *Proc. 2nd Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, Jun. 2018, pp. 1209–1214.

[30] S. Marne, S. Churi, and M. Marne, "Predicting breast cancer using effective classification with decision tree and k means clustering technique," in *Proc. Int. Conf. Emerg. Smart Comput. Informat. (ESCI)*, Mar. 2020, pp. 39–42.

[31] P. Kumar and S. K. Wasan, "Analysis of X-means and global K-means USING TUMOR classification," in *Proc. 2nd Int. Conf. Comput. Autom. Eng. (ICCAE)*, Feb. 2010, pp. 832–835.

[32] M. Umamaheswari and P. I. Devi, "Prediction of myocardial infarction using K-medoid clustering algorithm," in *Proc. IEEE Int. Conf. Intell. Techn. Control, Optim. Signal Process. (INCOS)*, Mar. 2017, pp. 1–6.

[33] S. Safdar, S. Ahmed Khan, A. Shaukat, and M. U. Akram, "Genetic algorithm based automatic out-patient experience management system (GAPEM) using RFIDs and sensors," *IEEE Access*, vol. 9, pp. 8961–8976, 2021.

[34] F. Khalid and A. N. Abbasi, "Challenges faced by pakistani healthcare system: Clinician's perspective," *J. College Phys. Surgeons Pakistan*, vol. 28, no. 12, pp. 899–901, 2018.

[35] A. Solanas, F. Casino, E. Batista, and R. Rallo, "Trends and challenges in smart healthcare research: A journey from data to wisdom," in *Proc. IEEE 3rd Int. Forum Res. Technol. Soc. Ind. (RTSI)*, Oct. 2017, pp. 1–6.

[36] A. Grocott and W. McSherry, "The patient experience: Informing practice through identification of meaningful communication from the patient's perspective," *Healthcare*, vol. 6, p. 26, Mar. 2018.

SARIA SAFDAR received the M.S. degree in computer software engineering from the National University of Science and Technology (NUST), Islamabad, Pakistan, in 2012, where she is currently pursuing the Ph.D. degree. She is a Lecturer with the Fatima Jinnah Women University (FJWU), Rawalpindi, Pakistan. Her research interests include health informatics, computer networks, data mining, and machine learning algorithms.

SHOAB AHMED KHAN received the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA. He is currently a Professor of computer and software engineering with the College of Electrical and Mechanical Engineering, National University of Sciences and Technology (NUST). He is an inventor of five awarded U.S. patents and has over 260 international publications. His book on digital design is published by John Wiley and Sons and is being followed in national and international universities. He has more than 22 years of industrial experience in companies in USA and Pakistan. He received Tamgh-e-Imtiaz (Civil), the Highest National Civil Award at Pakistan, the National Education Award, in 2001, and the NCR National Excellence Award in Engineering Education. He is the Founder of the Center for Advanced Studies in Engineering (CASE) and the Center for Advanced Research in Engineering (CARE). CASE is a primer engineering institution that runs one of the largest postgraduate engineering programs in the country and has already graduated 50 Ph.D. students and more than 1800 M.S. students in different disciplines in engineering, whereas CARE, under his leadership, has risen to be one of the most profound high technology engineering organizations in Pakistan developing critical technologies worth millions of dollars for organizations in Pakistan. CARE has made history by winning 13 PASHA ICT awards and 11 Asia–Pacific ICT Alliance Silver and Gold Merit Awards while competing with the best products from advanced countries, such as Australia, Singapore, Hong Kong, Malaysia. He has served as the Chairperson of the Pakistan Association of Software Houses (P@SHA) and as a member of the Board of Governance of many entities in the Ministry of IT and Commerce. He has also served as a member of the National Computing Council and the National Curriculum Review Committee.

ARSLAN SHAUKAT received the B.S. and M.S. degrees in computer engineering from the National University of Science and Technology (NUST), Islamabad, Pakistan, in 2003 and 2005, respectively, and the Ph.D. degree in computer science from The University of Manchester, U.K., in 2010. Currently, he is an Assistant Professor with the Department of Computer and Software Engineering, CEME, NUST. He has published various academic papers in refereed journals and conference proceedings. His research interests include machine learning, pattern recognition, digital image, and speech processing. He has been a member of technical program committees of numerous international conferences and also a reviewer of various IEEE journals and transactions. He was a recipient of academic awards including the Best Teacher Award, in 2018, and the Best Research Paper Award, in 2019. He jointly chaired the AI and IT convergence session at IEEE ITEC 2019 Conference held in South Korea.