This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3349953

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Smart Medical Appointment Scheduling: Optimization, Machine Learning and Overbooking to Enhance Resource Utilization

Catalina Valenzuela-Núñez [1], Guillermo Latorre-Núñez [2], Fredy Troncoso Espinosa [3].
[1]Department of Industrial Engineering, Universidad del Bío-Bío (e-mail: cvalenzuelan@ubiobio.cl)
[2]Department of Industrial Engineering, Universidad del Bío-Bío (e-mail: glatorre@ubiobio.cl)
[3]Department of Industrial Engineering, Universidad del Bío-Bío (e-mail: ftroncos@ubiobio.cl)

Corresponding author: Catalina Valenzuela-Núñez (e-mail: cvalenzuelan@ubiobio.cl)

ABSTRACT Scheduling medical appointments plays a fundamental role in managing patient flow and ensuring high-quality care. However, no-shows can significantly disrupt this process and affect patient care. To address this challenge, healthcare facilities can adopt different strategies, including overbooking in medical consultations. While this reduces the risk of unused slots, it can generate associated costs and affect the perception of service quality. In this article, we propose an integer linear optimization model that maximizes the expected utility of a medical center, considering the risk of no-shows and overbooking. For this purpose, machine learning is used to estimate the propensity of each patient to attend their medical appointment, using real data from three medical specialties of a hospital. The results of the application demonstrate the model's ability to assign appointments and perform overbooking efficiently and in an organized manner, implying an improvement in the utility of a medical center and a positive impact on the perception of the quality of care.

INDEX TERMS Scheduling, Medical Appointments, Overbooking, Machine Learning, Optimization, Healthcare.

I. INTRODUCTION

APPOINTMENT scheduling and booking for medical appointments play a crucial role in healthcare centers, as patients often require a variety of resources, such as doctors, nurses, equipment, and examination rooms, for their care. Patient waiting time is a key performance and quality indicator for hospitals, as excessive waiting can lead to reduced patient satisfaction and perception of the quality of care [1]. Thus, it becomes essential to efficiently utilize available resources to deliver high-quality services to healthcare system users [2].

One of the major management challenges in hospitals is the problem of patient appointment scheduling. This involves making decisions such as assigning physicians, exam rooms, and finding optimal times for patients to receive care [3]. However, it is equally important to consider the likelihood of patients not showing up for medical consultations, as their absence not only impacts their own health but also interferes with the care process for other patients in the facility. Understanding and proactively managing the possibility of patient no-shows plays a key role in achieving comprehensive optimization of hospital tasks, thereby elevating quality and patient-centered care.

Patients' health can be significantly impacted by increased indirect wait times resulting from patient no-shows. To develop an efficient and patient-friendly service, it is necessary to improve methods of prioritizing waiting times based on symptomatic experience, as highlighted by [4]. In countries with both public and private hospitals, longer waiting times are observed in the public sector, leading patients to seek faster care at private facilities [5]. These prolonged waiting times can be attributed to a variety of factors, such as the lack of effective techniques for medical appointment planning and scheduling [6].

In an effort to reduce the rate of nonattendance, healthcare centers commonly employ reminders or penalties. However, several studies indicate that these strategies only achieve a slight reduction in patient no-shows [7].

Therefore, one effective approach to optimize resource utilization is through the use of appointment scheduling systems, designed to obtain patient assignments that minimize measures, such as waiting time or physician idle time. These systems improve the utilization of costly medical resources, both personnel and facilities, while simultaneously reducing patient waiting times.

When considering medical appointment scheduling rules, two critical parameters come into play: the number of patients scheduled in each time slot and the duration of the time slot. The former determines the block size, i.e., the number of patients to be attended, while the latter represents the interval between visits [8]. There are various combinations of these parameters, but it is usual to assign a single patient to each time slot, varying the duration of their consultation.

Furthermore, the presence of late arrivals, no-shows, walk-ins, and emergencies can disrupt the assignment schedule. Physicians may also experience delays during clinical sessions or be interrupted by non-consulting activities [9]. All of these factors can significantly affect the scheduling and flow of visits in a healthcare environment.

The purpose of this research is to propose a model that allows incorporating overbooking in medical appointments, considering the propensity that a patient attends their medical appointment. We also use machine learning to determine this propensity and, for this reason, we present an objective function that represents the expected utility of the medical center.

The document is structured as follows: Section 2 discusses related work on medical appointment scheduling, covering concepts related to scheduling, machine learning, and problem-solving methods. Section 3 provides details about the materials and methods utilized in the study. Sections 4 and 5 present a comprehensive summary of the study's results. Finally, Section 6 discusses the conclusions drawn from the research and outlines potential avenues for future work.

II. RELATED WORK

Scheduling plays a fundamental part in resource allocation, whether it involves assigning personnel, determining material quantities for a process, or coordinating machinery. The efficiency of operations and service delivery is highly dependent on effective scheduling. In healthcare, there are several challenges to patient care, and since the 1950s, researchers have studied and proposed numerous models to address these issues [10]. These proposals aim to improve patient care through the development of new healthcare systems.

When it comes to the mathematical models used to solve the problem of scheduling medical appointments, many scheduling systems rely on heuristics, dynamic programming, and stochastic scheduling, as they are more resilient to factors such as random arrivals and service times. However, advancements in data science now allow for more accurate predictions of these factors, enabling the use of deterministic planning systems. Deterministic models are formulated using integer or mixed-integer linear programming to optimize specific performance measures for scheduling [8].

The phenomenon of patients not attending scheduled appointments is observed in healthcare facilities around the world. Missed appointments result in resource downtime, lower resource utilization, and reduced productivity. In contrast to the physical waiting time at the facility, there is an "indirect" waiting time between the referral date and the actual booking date, which can be crucial for the early diagnosis of disease. Patient no-shows not only impact organizational costs but also affect the effectiveness of healthcare services provided to these patients [11]. To mitigate the effects of nonattendance, overbooking is a popular strategy, as it allocates more capacity to the visit schedule, leading to more timely care.

Overbooking involves scheduling more than one patient for the same booking, similar to how airlines manage flight reservations. However, it can cause collisions when more than one person arrives for the same visit. In healthcare centers, patients cannot be simply "bumped" from their appointments as can happen in airlines. Consequently, such collisions can increase waiting times, affecting subsequent visits [12]. Usually, overbooking is done blindly, without considering the likelihood of a patient no-show. A more effective approach would involve overbooking a patient when it is highly likely that the initially scheduled patient will not show up.

A. SCHEDULING OF MEDICAL APPOINTMENT

The literature on the problem of patient appointment scheduling focuses on two crucial variables: the time interval between visits and the actual duration of the visit. The first variable signifies the scheduled duration, while the second represents the time a patient actually spends in their visit. Studies consider different time intervals for scheduling, assuming deterministic but unknown visit duration's.

In [13], the authors explore three medical appointment interval structures and the use of overbooking to address patient no-shows. Introducing flexibility in visit start times reduces waiting times while maintaining service provider efficiency, supported by simulation experiments [14]. An alternative approach to this problem is to adopt variable duration for medical appointments, as proposed by [15], [16], where the lengths of consultations depend on their start time.

The literature on this topic also considers patient priority, giving certain patients preferential treatment for early attention based on their medical condition or other relevant factors [8], [15], [17], [18]. Patient no-shows significantly impact scheduling systems. Assign-

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3349953

ing probabilities based on historical data helps gauge prediction sensitivity [13], [17].

B. MACHINE LEARNING IN MEDICAL ATTENDANCE

Supervised learning has been shown to be effective in predicting patient attendance at medical appointments [19]. Techniques such as linear regression, multiple regression, time series analysis, decision trees, and neural networks are commonly used for medical visit scheduling, optimizing it based on predefined parameters.

In [20], [21], machine learning techniques were employed to estimate the probability of patient no-shows, optimizing patient waiting time and physician overtime. Selecting the most accurate model for scheduling medical visits requires a systematic approach [22], [23].

Missed medical appointments can result in healthcare facilities not using resources optimally. Among the most common reasons cited by patients for not attending include forgetfulness and lack of communication with the healthcare facility [24]. Moreover, patients with emergency-related activities and postoperative care are more prone to missing appointments [25]. To tackle no-show rates, patient reminders via phone calls, emails, or text messages have been implemented. Consequently, healthcare organizations are turning to data science technologies to leverage available information.

A literature review of 50 articles conducted in 2020 explored the prediction of patient no-shows for medical consultations. Regression models emerged as the most frequently used technique to predict these absences [7]. Variables commonly included in machine learning research on medical appointment scheduling since 2017 encompass age, gender, day of the week, waiting time between consultation and appointment, previous no-show history, time of day, and distance from the healthcare facility.

C. SCHEDULING AND MACHINE LEARNING IN MEDICAL APPOINTMENTS

Researchers such as [8], [15], [16], [18], [20], [26], [27] consider various probabilities of patient no-shows using techniques such as machine learning, reference values from other studies, and process simulation. The purpose of having different predictions of patient attendance for medical visits is to use them as input data for the scheduling model to determine the optimal solution. [15], [20], [21], [26], [28] utilize machine learning methods to obtain probability values for patient attendance at medical appointments. These researchers apply techniques that classify user profiles based on the aforementioned variables in the machine learning section, associating them with a higher probability of no-shows. The objective is to propose strategies to reduce this no-show rate, such as reminders, appointment cancellations, and overbooking.

The benefit of working with machine learning and artificial intelligence is that it enables the extraction of information from the data, helping to uncover hidden patterns. It is also possible to develop a classification model that fits the data set using techniques that learn and predict the future attendance of patients at their medical appointments with a certain associated probability. Obtaining this value provides certainty that the scheduling model is functional for real-world application in the medical appointment problem. However, currently, there are relatively few research studies linking machine learning to appointment scheduling. On one hand, there is a need to separate patient groups classified as "attendees" and those with a higher probability of no-shows. On the other hand, there is a need to determine how and when to schedule these patients.

III. PROPOSAL OF A MEDICAL APPOINTMENT SCHEDULING MODEL WITH MACHINE LEARNING AND OVERBOOKING

In this section, we propose the integer linear programming model that allows the scheduling of medical appointments. The model will consider the following aspects:

- Three types of patients: the priority patients who are individuals with some degree of priority for medical care; the first-time patients who are first-time visitors to any medical specialty at the center. These have a higher priority for care compared to recurrent patients, but lower than those categorized as "priority patients". Finally, there are the regular patients, who have a history of more than one previous visit to the service and are not categorized as a priority for scheduling appointments.
- The main strategy of the medical center will be the use of overbooking to maximize the use of medical resources and some patients may require more than one care slot. The medical care slots have the same duration. To illustrate an example of this, Figure [1] show six patients scheduled in four time slots. Patient 4 needs two slots for his medical attention. Slots 2 and 3 are overbooked, with two and three patients respectively.

		Pacient 5 *No show*	
	Pacient 4 *Show*	Pacient 4 *Show*	
Pacient 1 *Show*	Pacient 2 *No show*	Pacient 3 *Show*	Pacient 6 *Show*
Slot 1	**Slot 2**	**Slot 3**	**Slot 4**

FIGURE 1. **Example of a schedule with overbooking in slots 2 and 3.**

- Each patient to be scheduled has a propensity to attend a medical appointment, which can also be interpreted as the probability of attendance (1-probability of non-attendance).

- There is a maximum number of patients to be overbooked per slot. This number is determined by a maximum value of the sum of their propensities to attend the medical appointment. Figure [1] shows that the maximum number of patients to be scheduled in the second slot is 2, and in the third slot is 3. Patients 2 and 5 have a low propensity to attend a medical appointment (shown in grey). When their propensity are combined with those of others who have a higher propensity, they do not reach the established limit.
- The primary objective of the medical center is to optimize its utility which depends mainly on the number of patients scheduled, the propensity to attend of each patient and the quantity of slots required to provide care for each patient.

Next, the utility function of a medical center is obtained, and then the optimization model is proposed.

A. FORMULATION OF THE OBJECTIVE FUNCTION

Before introducing the integer linear programming model, we present an expected utility function for a medical center in terms of the utilization of care slots, as shown in Equation (1). This equation provides the criterion by which the model will make patient scheduling decisions over a given time horizon.

$$\mathrm{U} = \sum_{\mathrm{P}} \mathrm{IPrb_p SP_p} - \mathrm{C(k)} \ \forall \mathrm{p} \in \mathrm{P} \tag{1}$$

The first component of the utility function represents the expected revenue. $\mathbf{I}$ represents the revenue for using a time slot, and $\mathbf{P}\mathrm{rb_p}$ is the propensity of a patient p to show up for his or her medical appointment. To obtain this value, we consider a set of attributes related to the patient's behavior in previous medical appointments. This value is given by (2), in which $\mathbf{A}_\mathrm{p}$ is the set of relevant attributes of patient p, and $\mathbf{f}$ is a function, chosen under a certain context, that transforms A_p into a propensity value.

$$\mathrm{Prb_p} = \mathrm{f(A_p)} \ \forall \mathrm{p} \in \mathrm{P} \tag{2}$$

$\mathbf{SP_p}$ is the number of slots required to attend to patient p. The second term represents the cost of overbooking. This cost represents all costs incurred due to the delay in patient care and the cost of extra care box time. These costs are represented by $\mathbf{C(k)}$ and depend on the number k of overbooked patients in a care slot. We consider the cost $C(k)$ as a fraction of the revenue I. We will denote this fraction $\alpha_\mathbf{k}$, which will depend on the number k of overbooked patients. Thus, the utility function is as shown below in (3).

$$\mathrm{U} = \sum_{\mathrm{P}} \mathrm{IPrb_p SP_p} - \mathrm{I}\alpha_\mathrm{k} \ \forall \mathrm{p} \in \mathrm{P} \tag{3}$$

In this investigation $\mathbf{Prb_p}$ will be obtained by using machine learning.

B. MATHEMATICAL FORMULATION OF THE MODEL

In the mathematical formulation of the model, we use the expected utility function (3) as the objective function. Since revenue I is present in both terms, we set $I = 1$ and interpret α_k as the penalty for overbooking.

We use the following notation and definitions. Let $\theta = \{\forall\ \mathrm{p} \in \mathrm{P}, \mathrm{d} \in \mathrm{D}, \mathrm{b} \in \mathrm{B}, \mathrm{s} \in \mathrm{S_{db}} : \mathrm{s} + \mathrm{SP_p} - 1 \leq \mathrm{S_{db}},\ \mathrm{S_{db}} > 0\}$ represent the notation used in the constraints. Furthermore, it is important to clarify that ss refers to a slot different from s but still belonging to the set $\mathbf{S}$, which starts in $\mathrm{s} + \mathrm{SP_p} - 1$. The set $\mathrm{O} = \{(\mathrm{d}, \mathrm{b}) : \ \forall\ \mathrm{d} \in \mathrm{D}, \mathrm{b} \in \mathrm{B}\}$ corresponds to the set of available time slots for overbooking.

The proposed model for scheduling medical appointments is as follows:

<u>Sets</u>

$\mathbf{P}$ = set of patients to be scheduled,
$\mathbf{D}$ = set of days to be scheduled,
$\mathbf{B}$ = set of care box,
$\mathbf{O}$ = set of slots of the box b on day d where overbooking is allowed,
$\mathbf{S_{db}}$ = set of slots available on d in the box b.

<u>Parameters</u>

$\mathbf{q_d}$ = proportion of slots available for first-time patients on day d,
$\mathbf{f_p}$ = binary parameter that is equivalent to 1 when patient p has a priority for attention and 0 otherwise,
$\mathbf{G_p}$ = binary parameter that is equivalent to 1 when patient p corresponds to first visit and 0 otherwise,
$\mathbf{Prb_p}$ = propensity that patient p will show up for his medical appointment,
$\mathbf{SP_p}$ = number of slots needed to attend the patient p,
$\alpha_\mathbf{k}$ = overbooking penalty,
$\mathbf{k}$ = maximum number of patients assigned to the same slot (overbooking),
$\mathbf{maxPrb}$ = maximum likelihood sum for patients assigned to the same slot.

<u>Decision variables</u>

$\mathbf{x_{pdbs}}$ = binary variable that is equal to 1 if patient p is assigned to slot s on day d in box b and 0 otherwise,
$\mathbf{y_p}$ = binary variable that is equal to 1 if patient p is programmed and 0 otherwise,
$\mathbf{u_{pdbs}}$ = binary variable that is equal to 1 if the patient p uses the slot s on day d in the box b and 0 otherwise,
$\mathbf{B_{dbs}}$ = binary variable that is equal to 1 if there is overbooking in the slot s in the box b on day d and 0 otherwise.

Objective function

$$\mathrm{Max Z} = \sum_{\theta} \mathrm{y_p Prb_p SP_p} - \sum_{\theta} \alpha_k \mathrm{B_{dbs}} \tag{4}$$

subject to:

$$\mathrm{y_p} = \sum_{\theta} \mathrm{x_{pdbs}} \;\; \forall \;\; \mathrm{p, d, b, s} \in \theta \tag{5}$$

$$\sum_{\theta} \mathrm{u_{pdbs}} = \mathrm{SP_p y_p} \;\; \forall \;\; \mathrm{p, d, b, s} \in \theta \tag{6}$$

$$\mathrm{x_{pdbs} SP_p} \leq \sum_{\theta} \mathrm{u_{pdbss}} \sum_{\mathrm{ss \in s + SP_p - 1}} \mathrm{u_{pdbs}} \qquad \forall \;\; \mathrm{p, d, b, s} \in \theta \tag{7}$$

$$\sum_{\theta} \mathrm{u_{pdbs}} \leq 1 \;\; \forall \;\; \mathrm{p, d, b, s} \in \theta \tag{8}$$

$$\sum_{\theta} \mathrm{u_{pdbs}} \leq 1 + \mathrm{B_{dbs}(maxp} - 1) \qquad \forall \;\; \mathrm{p, d, b, s} \in \theta \;\wedge\; \forall \; \mathrm{(d, b)} \in \mathrm{O} \tag{9}$$

$$\sum_{\theta} \mathrm{x_{pdbs}} \leq 1 \;\; \forall \;\; \mathrm{p, d, b, s} \in \theta \tag{10}$$

$$\sum_{\theta} \mathrm{x_{pdbs} G_p SP_p} \geq \min\Big(\sum_{\mathrm{p \in P}} \mathrm{G_p SP_p}, \; \mathrm{q_d} \sum_{\mathrm{(d,b) \in O}} \mathrm{S_{db}}\Big) \qquad \forall \;\; \mathrm{p, d, b, s} \in \theta \tag{11}$$

$$\sum_{\theta} \mathrm{x_{pdbs} f_p SP_p} \geq \min\Big(\sum_{\mathrm{b \in B}} \mathrm{S_{db}} - \mathrm{q_d} \sum_{\mathrm{b \in B}} \mathrm{S_{db}}, \sum_{\mathrm{p \in P}} \mathrm{f_p SP_p}\Big) \qquad \forall \;\; \mathrm{p, d, b, s} \in \theta \tag{12}$$

$$\sum_{\theta} \mathrm{u_{pdbs}} \geq 1 \;\; \forall \;\; \mathrm{p, d, b, s} \in \theta \tag{13}$$

$$\mathrm{x_{pdbs}, y_p, u_{pdbs}, B_{dbs}} \in 0, 1 \qquad \forall \;\; \mathrm{p, d, b, s} \in \theta \;\wedge\; \forall \; \mathrm{(d, b)} \in \mathrm{O} \tag{14}$$

The model starts by scheduling priority patients with the parameter $f_p = 1$, followed by first-time patients represented by $G_p = 1$, and finally, regular patients represented by $G_p = 0$. All this while ensuring that all available slots for each day of the planning horizon are occupied. Besides, each patient requires a certain number of slots for their care, represented by the parameter SP_p. Generally, it is assumed that each patient will require one slot ($SP_p = 1$), while priority patients are assigned two slots ($SP_p = 2$), considering that they may need more extensive attention.

According to the model, each patient is scheduled for a specific day with the allocation of the necessary slots for his or her appointment. It is important to note that only one patient can be seen at a time. In addition, a maximum limit of 120% (**maxPrb**) is considered for the sum of the probabilities of the scheduled patients to avoid overbooking. In other words, if a patient has a 70% probability of attending, the model will overbook another patient who has, at most, a 50% probability of attending. This ensures that the sum of the probabilities does not exceed the 120% limit set for each scheduled slot.

Finally, it is established that the medical care center is allocated a minimum number of exclusive slots for first-time patients and priority patients. If, after assigning these patients to their respective slots, there are still slots available, then they are assigned to regular patients.

The objective function maximizes the probability of slot occupancy of the medical center considering the penalty for overbooking. Constraint (5) states that if patient p is scheduled, they must be assigned a day d, a slot s, and a box b. Constraint (6) indicates that if patient p is scheduled, the number of assigned slots must match their requirements. Constraint (7) ensures that the slots assigned to patient p are exclusively theirs. Constraint (8) prevents assigning more than one patient to a non-overbooking slot. Constraint (9) captures the slots where multiple patients are assigned and limits the maximum number of patients assigned to the same slot. Constraint (10) ensures that a patient can only be scheduled once. Constraints (11) and (12) establish a minimum number of slots to be assigned to first-time patients and prioritized patients, respectively. Lastly, constraint (13) requires all available slots to be occupied, and constraint (14) ensures non-negativity of the variables.

IV. MODEL APPLICATION TO REALWORLD DATASET

In this section, the model proposed above will be applied to real data obtained from the medical appointment history of three medical specialties at Dr. Guillermo Grant Benavante Hospital in the city of Concepción, Chile. The database contains medical consultation records from the first semester of the year 2021, with a total of approximately 340,000 entries. The three selected medical specialties were those with a monthly non-attendance rate greater than 10%. These specialties are Neurology, Gynecology, and Otorhinolaryngology. Consequently, the database comprises 10,000 medical appointment records.

The data presented in this study is available upon request to the respective hospital. Due to its private and sensitive nature, which contains personal and confidential information about the patients, the data is not publicly accessible.

An important aspect for the model's application is determining the propensity of each patient to attend a medical appointment, as shown in Equation (2). Given that we have historical data related to the patients' attendance at their previous medical visits, machine learning will be used to identify patterns to determine this propensity.

This section provides details on how patient attendance propensity has been calculated using machine learning techniques. A machine learning algorithm is utilized to identify specific characteristics of patients who are more likely to miss their medical appointments. Test instances are also created to verify the functionality of the model

A. DETERMINATION OF THE PROPENSITY TO ATTEND MEDICAL APPOINTMENTS BY MACHINE LEARNING

Regarding Prb_p (Equation (2)), in this investigation, the function $f(;)$ will be a Machine Learning algorithm. This trained algorithm will allow us to obtain the behavior pattern of patients based on a set of A_p attributes that describe the historical background related to the patients' attendance at appointments. Through the application of the trained algorithm, we will obtain the propensity to attend their medical appointment for each patient to be scheduled.

The historical background of patients in the database included attributes such as the date of the medical appointment, type of care, medical specialty, patient's age, commune of residence of the patient, among others. After a data cleaning process, it was necessary to apply the SMOTE algorithm [29] to generate a balance between the number of attendees and absentees in the database

The most relevant attributes were determined using the information gain criterion [30]. These include the difference in days between the creation of the record and the actual date of the medical appointment, the type of activity and medical care required by the patient, the number of times the patient has previously missed and attended medical appointments, and the establishment of origin from which these patients come, among others.

After identifying these most relevant attributes, five machine learning models were trained and tested using Rapidminer software [31]: Decision Tree [32], Neural Network [33], Support Vector Machine [34], Linear Regression [35], and Naive Bayes [36].

For training, testing, and selection of the most appropriate machine learning algorithm, the process shown in Figure [2] was used. The training process consists of a training subset and a validation subset. Using data from January to April 2021 as the training set, the algorithm was trained to learn the attendance patterns of the patients present in the data. Subsequently, using data from May 2021 for validation, the parameters of each algorithm were adjusted to achieve the best possible predictions for this month with each of them. After parameter adjustment, the performance of each model was tested by predicting the month of June 2021. The results of this testing process are presented in Table [1].

Table [1] shows the predictive performance of each machine learning algorithm using four hit-based indicators. Accuracy represents the overall proportion of patients

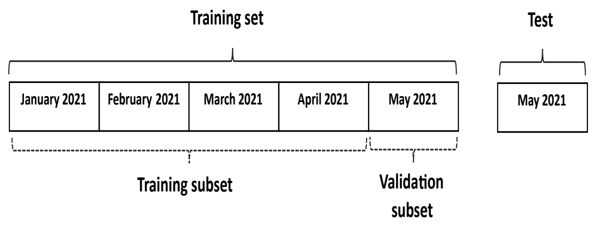

FIGURE 2. Training and validation set for the algorithms.

TABLE 1. Results of machine learning algorithms.

Algorithm	Accuracy	Precision	Recall	AUC
Decision Tree	0.745	0.758	0.720	0.800
Neural Net	0.758	0.772	0.733	0.828
SVM	0.726	0.659	0.937	0.851
Logistic Regression	0.674	0.698	0.614	0.734
Naive Bayes	0.637	0.665	0.550	0.701

correctly predicted by the model. Precision is the number of patients correctly predicted to "attend" divided by the total number of patients the algorithm predicts will "attend". Recall is the number of patients correctly predicted to "assist" divided by the total number of patients who actually showed up. AUC (Area Under the Curve) is an indicator that measures the area under the ROC [37], [38] curve. The ROC curve is a graphical representation of the performance of a classifier that relates the true positive and false positive rates in a binary prediction classifier performance. The value of each indicator varies between 0 and 1. The higher the value, the better the performance of the model according to the indicator criteria. The Support Vector Machine (SVM) algorithm is chosen, as it demonstrates better performance in the Recall and AUC indicators.

Figure [3] shows the methodology to apply a machine learning algorithm to predict medical appointment attendance. It begins with loading historical patient attendance and no-show data for a specified period of time. Next, the relevant variables for the model are extracted, the algorithm is applied and the results are used to implement strategies for scheduling medical appointments, in our case, an integer linear programming model that considers overbooking. The results delivered by the algorithm is a value between 0 and 1 where a higher value represents a higher propensity to show up for an medical visit.

B. SCHEDULING MEDICAL APPOINTMENTS USING THE MODEL

After determining the propensity, we employ the proposed model to create a weekly schedule. To do this, we initially assume that each time slot has a duration of 15 minutes, aligning with the average time it takes for each patient to receive medical attention. Consequently, on average, each patient will occupy only one time slot for their scheduled procedure, except in cases where

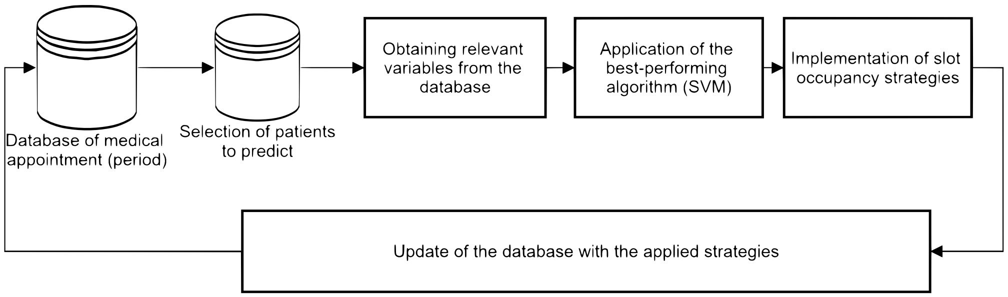

FIGURE 3. Diagram of the process of identification and countermeasures for patients with absenteeism.

two time slots are required. Additionally, it has been established that doctors maintain a continuous schedule of 8 hours, starting from 8:00 A.M. and concluding at 4:00 P.M., equivalent to 32 time slots of 15 minutes each. It is important to note that the number of consultation rooms may vary across specialties; however, for the purpose of observing patient scheduling behavior, we will consider only one consultation room (box) as a reference.

We estimate $k = 2$ to mitigate inefficiencies arising from overcrowding and longer waiting times. Given this value of $k = 2$, it was necessary to calculate the value of $\alpha_{\mathrm{k=2}}$. To achieve this, 9 groups of patients were scheduled, and for each group, the alpha parameter was varied from 0.1 to 0.9 in increments of 0.1. The sizes of the 9 patient groups were 200, 250, 300, 350, 400, 450, 500, 550, and 600. For each alpha value and group size, the occupancy rate was calculated, defined in the Equation (15), where the denominator corresponds to the occupancy propensity of the slots and the nominator is the number of slots scheduled for the patients.

$$\text{Occupancy rate} = \frac{\sum \mathrm{y_p Prb_p SP_p}}{\sum \mathrm{y_p SP_p}} \tag{15}$$

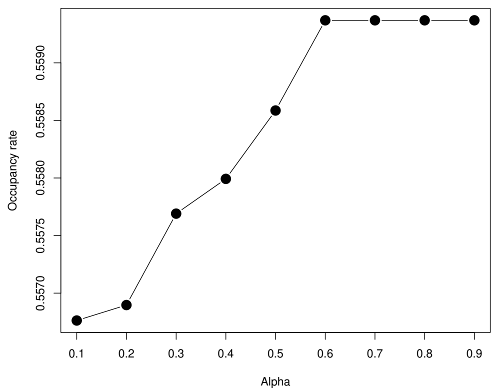

FIGURE 4. Occupancy rate based on $\alpha_{k=2}$ value.

The average occupancy rate obtained for each $\alpha_{\mathrm{k=2}}$ value is shown on the Figure 4. For low values of alpha (0.1, 0.2, and 0.3), the model employs a higher degree of overbooking in patient scheduling. For alpha values of $\alpha_{\mathrm{k=2}} = 0.4$ and 0.5, the model reduces the extent of overbooking. Starting from an alpha value of $\alpha_{\mathrm{k=2}} = 0.6$, the model no longer utilizes overbooking, resulting in suboptimal outcomes. The highest occupancy rate is attained with an alpha value of $\alpha_{\mathrm{k=2}} = 0.5$, and this value is selected (the literature suggests $\alpha_{\mathrm{k=2}} = 0.3$ [15]).

Tables [2], [3] and [4] show the results obtained from the application of the model to different groups of patients, with values of k = 2 and $\alpha_{\mathrm{k=2}} = 0.5$, illustrating the occupancy rate trends for each medical specialty. The first column $\sum \mathrm{y_p}$ refers to the number of patients scheduled by the model, the second column $\sum \mathrm{y_p SP_p}$ is the slots scheduled for patients, the third column $\sum \mathrm{y_p SP_p Prb_p}$ is the occupancy propensity of these slots, the fourth column is the occupancy rate and then there is the value of the objective function (Z) and the value of the overbooking. From this it can be seen that the three specialties achieve an occupancy rate of up to 57%. However, this value varies depending on the size of the patient cohort.

C. EFFECTIVENESS OF THE APPLICATION OF THE MODEL

To assess the model's effectiveness, the scheduling approach will be compared to the current hospital practices, which involve random scheduling but prioritize patients defined as high-priority and those attending for their first visit.

Tables [5], [6], and [7] display the resulting occupancy rates for each patient group when scheduled according to the hospital's current practices.

It is observed that the values vary depending on the size of the group to be scheduled and the propensities of the scheduled patients to attend their medical appointments.

Figure [5] illustrates the comparison of the average occupancy rates obtained using the proposed model (Model) and the current hospital practices (Random). The average occupancy rate is calculated as the mean of the occupancy rates obtained for each group of scheduled patients. Also demonstrates that at a confidence level of 0.95, the confidence intervals of the means do not

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3349953

TABLE 2. Results of appointment scheduling for the neurology specialty using the model.

N° of Patients	$\sum y_p$	$\sum y_p SP_p$	$\sum y_p Prb_p SP_p$	Occupancy rate	Z	$\alpha_k B_{dbs}$
200	153	190	102.26	0.5382	87.26	15.0
250	160	199	108.27	0.5441	88.77	19.5
300	160	199	109.6	0.5508	90.10	19.5
350	160	200	110.45	0.5523	90.45	20.0
400	160	196	109.72	0.5598	91.72	18.0
450	160	197	111.58	0.5664	93.08	18.5
500	160	196	111.85	0.5707	93.85	18.0
550	160	195	111.64	0.5725	94.14	17.5
600	160	196	113.31	0.5781	95.31	18.0

TABLE 3. Results of appointment scheduling for the gynecology specialty using the model.

N° of Patients	$\sum y_p$	$\sum y_p SP_p$	$\sum y_p Prb_p SP_p$	Occupancy rate	Z	$\alpha_k B_{dbs}$
200	158	192	105	0.5483	89.28	16.0
250	160	201	111	0.5513	90.31	20.5
300	160	200	112	0.5578	91.55	20.0
350	160	199	111	0.5602	91.97	19.5
400	160	201	113	0.5608	92.23	20.5
450	160	200	113	0.5661	93.22	20.0
500	160	198	112	0.5653	92.93	19.0
550	160	198	113	0.5707	94.00	19.0
600	160	200	115	0.5743	94.85	20.0

TABLE 4. Results of appointment scheduling for the specialty of otorhinolaryngology using the model.

N° of Patients	$\sum y_p$	$\sum y_p SP_p$	$\sum y_p Prb_p SP_p$	Occupancy rate	Z	$\alpha_k B_{dbs}$
200	160	200	111.09	0.5555	91.09	20.0
250	160	201	112.68	0.5606	92.18	20.5
300	160	199	112.04	0.5630	92.54	19.5
350	160	202	114.35	0.5661	93.35	21.0
400	160	203	115.63	0.5696	94.13	21.5
450	160	202	114.75	0.5681	93.75	21.0
500	160	200	114.43	0.5722	94.43	20.0
550	160	201	115.55	0.5749	95.05	20.5
600	160	203	117.67	0.5797	96.17	21.5

TABLE 5. Random scheduling (hospital) of appointments for the neurology specialty.

N° of Patients	$\sum y_p$	$\sum y_p SP_p$	$\sum y_p Prb_p SP_p$	Occupancy rate
200	135	160	91.83	0.5739
250	130	155	81.81	0.5278
300	140	160	88.25	0.5516
350	130	160	82.91	0.5182
400	120	160	86.26	0.5392
450	150	160	80.45	0.5028
500	135	160	76.92	0.4808
550	130	160	79.16	0.4948
600	140	160	80.10	0.5006

overlap. Consequently, a significant difference is evident between the average results generated by the model and the outcomes of the hospital's scheduling. These distinctions highlight that the model allows for more efficient scheduling of medical appointments by considering each patient's propensity to attend and implementing overbooking.

Figure [6] depicts the average performance of both the model and the hospital's scheduling. It is evident that the model achieves a significantly higher occupancy rate than the hospital's scheduling. This is supported by the non-overlapping confidence intervals of the means at a confidence level of 0.95.

To demonstrate that the average performance of the model is significantly higher, a mean test is conducted between the performance of both schedules. Before defining the appropriate test, it is necessary to check the normality of the data and the homogeneity of variance. We utilize the Shapiro-Wilk test for normality and the Bartlett test for variance homogeneity.

Table [8] displays the results of both tests. Since the p-values in the Shapiro-Wilk test for both Hospital and Model are greater than the significance value of 0.05, it can be concluded that there is no statistical evidence to reject the null hypothesis, indicating that the data follows a normal distribution. However, as the

TABLE 6. **Random scheduling (hospital) of appointments for the gynecology specialty.**

N° of Patients	$\sum y_p$	$\sum y_p SP_p$	$\sum y_p Prb_p SP_p$	Occupancy rate
200	125	155	89.89	0.5799
250	130	160	82.24	0.5140
300	125	160	84.04	0.5252
350	125	155	85.37	0.5508
400	125	160	82.94	0.5184
450	130	155	86.55	0.5584
500	135	160	81.42	0.5089
550	150	160	84.44	0.5277
600	125	160	83.34	0.5209

TABLE 7. **Random scheduling (hospital) of appointments for the otorhinolaryngology specialty.**

N° of Patients	$\sum y_p$	$\sum y_p SP_p$	$\sum y_p Prb_p SP_p$	Occupancy rate
200	120	160	86.88	0.5430
250	120	160	75.31	0.4707
300	110	160	81.07	0.5067
350	140	160	89.04	0.5565
400	120	160	75.73	0.4733
450	125	160	81.68	0.5105
500	130	160	81.59	0.5100
550	150	160	84.44	0.5278
600	120	155	80.28	0.5179

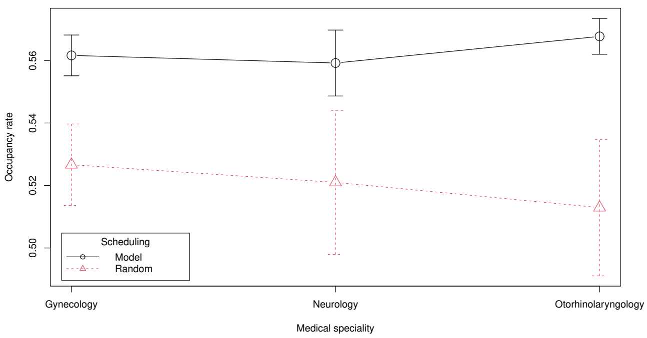

FIGURE 5. **Occupancy rates for medical specialties scheduled by model and hospital (random)**

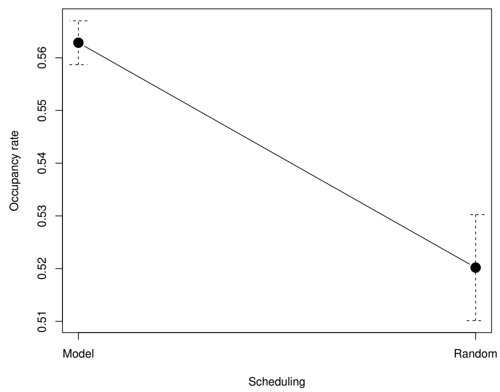

FIGURE 6. **Mean occupancy rate of the model and the hospital (random).**

p-value in the Bartlett test is less than the significance value of 0.05, there is statistical evidence to reject the null hypothesis, suggesting that the variances in both schedules are different.

Given that the data follows a normal distribution and there are significant differences in the variances between the two schedules, the most appropriate method to test the equality between the average performances of both schedules is the "Analysis of Variance with Heterogeneity of Variances," also known as Welch's ANOVA. Table [8] reveals that the p-value of this test is less than the significance value of 0.05, indicating a highly significant difference in the average occupancy rates of both schedules. This implies that using the proposed model for scheduling medical appointments achieves a higher occupancy rate than the current hospital practices.

TABLE 8. **P-value results of applied tests.**

Test	p-value
Shapiro-Wilk Hospital	0.7786
Shapiro-Wilk Model	0.5776
Bartlett	0.00002526
Welch's ANOVA	0.00000000189

D. COMPUTATIONAL RESULTS

The model was implemented using the AMPL language and solved using the CPLEX 22.1.1.0 solver. The solver was allocated a computational time limit of 3600 seconds, and the system specifications included an 11th Gen Intel(R) Core(TM) i7-11800H 2.30 GHz processor with 16.00 GB of RAM.

The computational results presented in Table [9] highlight the outstanding performance of the AMPL model.

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3349953

Each row corresponds to a specific medical specialty and provides information on the goodness-of-fit (GAP) and computation time in seconds for each instance. The model achieved an optimal solution with a GAP of 0.00%, indicating highly accurate solutions. Computation times varied based on the medical specialty and instance size. This efficiency allows the model to handle longer time horizons for appointment scheduling, making it a valuable tool for healthcare facilities seeking optimal and timely solutions.

TABLE 9. Computational results.

Medical specialty	GAP %	Comp. average time (sec)
Neurology	0.00%	477.71
Gynecology	0.00%	724.84
Otorhinolaryngology	0.00%	770.33

V. DISCUSSION

The proposed model aims to optimize the utilization of health center time slots by combining overbooking and considering a patient's propensity to attend their medical appointment. It prioritizes first-time patients, followed by high-priority and recurrent patients, ensuring comprehensive coverage of all available slots for the day. The model incorporates various constraints based on facility policies and can be tailored to specific scenarios.

The test instances demonstrated the model's ability to logically and efficiently schedule patients, resulting in a high number of scheduled patients and overbooked slots. Tables [2], [3], and [4] show that the number of scheduled patients $\sum y_p SP_p$ is consistently greater than the number of available slots $\sum y_p$. It is also possible to observe that the higher the number of patients to be scheduled, the higher the value of Z and the higher the occupancy rate. This is due to the fact that by increasing the universe of patients to be scheduled, the model can choose a subset with a higher propensity.

Occupancy rates in all cases indicate the model's effectiveness in efficiently allocating patient appointments to available time slots, resulting in a well-organized and efficient scheduling system. This is evident in the higher average occupancy rate provided by the model in each specialty (Figure [5]) and in the overall average occupancy rate (Figure [6]) compared to the hospital's scheduling occupancy rate.

The model achieves an average occupancy rate of 56.29%, which is significantly higher than the 52.26% achieved by the hospital. This level of occupancy is considered high, as it guarantees that at least 56% of the scheduled slots for both instances will be occupied under any circumstances. This indicates that the scheduling strategy implemented by the model effectively fills a substantial portion of the available slots with patient appointments that have a higher propensity to attend, resulting in optimal resource utilization.

This is particularly crucial in healthcare environments as it ensures the optimal utilization of medical resources, leading to reduced patient waiting times and an overall enhancement in the quality of care provided.

It is important to note that the occupancy rate can be improved by increasing patients' attendance propensities. In this study, the average patient attendance rate is approximately 48%. Therefore, the model combines these attendance probabilities with the established parameters. For instance, when the model was applied to a set of patients with attendance probabilities exceeding 80%, while keeping the same parameters, the results demonstrated an occupancy rate exceeding 79%. This illustrates that slot occupancy is directly influenced by the probability of patient attendance.

The available information for estimating patients' propensity to attend their medical appointment was limited in comparison to other databases utilized in prior studies. These alternative datasets encompassed a more comprehensive set of patient characteristics, including factors such as waiting time, distance to the health facility, indigenous identity, life insurance status, educational level, and more.

The incorporation of additional data has the potential to yield more precise estimates of attendance probability, given that the applied machine learning technique is sensitive to the available information. Consequently, having access to a broader range of patient-related information enhances the accuracy of propensity to attend estimation.

VI. CONCLUSIONS AND FUTURE WORK

The implementation of an overbooking strategy allows for more efficient resource utilization; however, it is not without associated costs, such as increased waiting times for some patients, which directly impacts the perception of service quality.

The proposed integer linear optimization model incorporates the propensity of each scheduled patient to attend, in conjunction with the overbooking strategy. By maximizing the expected utility function of a medical center, this model demonstrates an efficient approach to scheduling medical appointments, ultimately improving the utility of a medical center.

For estimating the propensity to attend for each patient, the Support Vector Machine (SVM) machine learning algorithm exhibited superior performance, achieving an AUC of 85%. The utilization of SVM and its continuous retraining with new patient visit data positions the proposed model as an adaptive and intelligent medical appointment scheduling tool. The combination of optimization techniques and data-driven decision-making holds the potential to revolutionize medical appointment scheduling, enhancing access to care and overall healthcare service delivery.

The results from the application to real data from three medical specialties lead to the conclusion that the model efficiently allocates patients and care hours, minimizing the generation of unused slots. Despite the low propensity of patients in these three specialties to attend medical appointments and the prioritization of a substantial group of patients, the model demonstrated statistically superior results compared to the random scheduling employed by the hospital, while considering the same conditions of patient propensity and priority.

To further enhance the model's performance and robustness, the study will explore a utility function where the cost component is a function of the patients' propensity to attend their medical appointment.

Regarding the model's applicability, it is imperative to develop an integration with the data flow so that the SVM algorithm can be automatically retrained periodically. This will enable the propensity to attend to be updated in accordance with the dynamic behavior of the patients.

REFERENCES

[1] J. Sun, Q. Lin, P. Zhao, Q. Zhang, K. Xu, H. Chen, C. J. Hu, M. Stuntz, H. Li, and Y. Liu, "Reducing waiting time and raising outpatient satisfaction in a chinese public tertiary general hospital-an interrupted time series study," BMC Public Health, vol. 17, pp. 1–11, 2017.
[2] J. P. Sepúlveda and C. Berroeta, "Reglas de decisión para la selección de sistemas de citas basadas en características del paciente y de la institución de salud," Revista médica de Chile, vol. 140, pp. 867–872, 2012.
[3] G. Lamé, O. Jouini, and J. S.-L. Cardinal, "Outpatient chemotherapy planning: A literature review with insights from a case study," IIE Transactions on Healthcare Systems Engineering, vol. 6, pp. 127–139, 2016.
[4] S. Derrett, C. Paul, and J. M. Morris, "Waiting for elective surgery: effects on health-related quality of life," International Journal for Quality in Health Care, vol. 11, pp. 47–57, 1999.
[5] S. J. Duckett, "Private care and public waiting," Australian Health Review, vol. 29, pp. 87–93, 2005.
[6] M. Heshmat, K. Nakata, and A. Eltawil, "Solving the patient appointment scheduling problem in outpatient chemotherapy clinics using clustering and mathematical programming," Computers and Industrial Engineering, vol. 124, pp. 347–358, 2018.
[7] J. F. R. Saldaña, R. G. Flores, H. B. Harvey, C. Liu, J. Ai, C. Jaworsky, C. E. Guerrier, E. Flores, O. Pianykh, D. Carreras-García, D. Delgado-Gómez, F. Llorente-Fernández, and A. Arribas-Gil, "Predicting no-shows in radiology using regression modeling of data available in the electronic medical record," Ingenierías, vol. 8, pp. 37–47, 2017.
[8] D. Carreras-García, D. Delgado-Gómez, E. Baca-García, and A. Artés-Rodriguez, "A probabilistic patient scheduling model with time variable slots," Computational and Mathematical Methods in Medicine, vol. 2020, 2020.
[9] T. Cayirli and E. Veral, "Outpatient scheduling in health care: a review of literature," Production and operations management, vol. 12, pp. 519–549, 2003.
[10] R. Alizadeh, J. Rezaeian, M. Abedi, and R. Chiong, "A modified genetic algorithm for non-emergency outpatient appointment scheduling with highly demanded medical services considering patient priorities," Computers and Industrial Engineering, vol. 139, p. 106106, 2020.
[11] M. C. Rossi and H. Balasubramanian, "Panel size, office visits, and care coordination events: a new workload estimation methodology based on patient longitudinal event histories," MDM Policy and Practice, vol. 3, p. 2381468318787188, 2018.
[12] Y. Huang and D. A. Hanauer, "Patient no-show predictive model development using multiple data sources for an effective overbooking approach," Applied clinical informatics, vol. 5, pp. 836–860, 2014.
[13] Y. Chen, Y.-H. Kuo, P. Fan, and H. Balasubramanian, "Appointment overbooking with different time slot structures," Computers and Industrial Engineering, vol. 124, pp. 237–248, 2018.
[14] L. R. LaGanga and S. R. Lawrence, "Clinic overbooking to improve patient access and increase provider productivity," Decision Sciences, vol. 38, pp. 251–276, 2007.
[15] D. Ruiz-Hernández, D. García-Heredia, D. Delgado-Gómez, and E. Baca-García, "A probabilistic patient scheduling model for reducing the number of no-shows," Journal of the Operational Research Society, vol. 71, pp. 1102–1112, 2020.
[16] Y.-H. Kuo, H. Balasubramanian, and Y. Chen, "Medical appointment overbooking and optimal scheduling: tradeoffs between schedule efficiency and accessibility to service," Flexible Services and Manufacturing Journal, vol. 32, pp. 72–101, 2020.
[17] C. Zacharias and M. Pinedo, "Appointment scheduling with no-shows and overbooking," Production and Operations Management, vol. 23, pp. 788–801, 2014.
[18] M. Savelsbergh and K. Smilowitz, "Stratified patient appointment scheduling for mobile community-based chronic disease management programs," IIE Transactions on Healthcare Systems Engineering, vol. 6, pp. 65–78, 2016.
[19] A. A. R. Lozano, "Analítica de datos para la optimización del agendamiento de citas médicas," Expeditio Repositorio Institucional UJTL, 2020.
[20] S. L. Harris and M. Samorani, "On selecting a probabilistic classifier for appointment no-show prediction," Decision Support Systems, vol. 142, p. 113472, 2021.
[21] M. Samorani and S. Harris, "The impact of probabilistic classifiers on appointment scheduling with no-shows," in Fortieth International Conference on Information Systems, 2019.
[22] F. Rustam, A. A. Reshi, A. Mehmood, S. Ullah, B.-W. On, W. Aslam, and G. S. Choi, "Covid-19 future forecasting using supervised machine learning models," IEEE Access, vol. 8, pp. 101489–101499, 2020.
[23] J. Qin, L. Chen, Y. Liu, C. Liu, C. Feng, and B. Chen, "A machine learning methodology for diagnosing chronic kidney disease," IEEE Access, vol. 8, pp. 20991–21002, 2020.
[24] E. Kaplan-Lewis and S. Percac-Lima, "No-show to primary care appointments: why patients do not come," Journal of primary care and community health, vol. 4, pp. 251–255, 2013.
[25] A. S. Hwang, S. J. Atlas, P. Cronin, J. M. Ashburner, S. J. Shah, W. He, and C. S. Hong, "Appointment "no-shows" are an independent predictor of subsequent quality of care and resource utilization outcomes," Journal of general internal medicine, vol. 30, pp. 1426–1433, 2015.
[26] J. Lin, K. Muthuraman, and M. Lawley, "Optimal and approximate algorithms for sequential clinical scheduling with no-shows," IIE Transactions on Healthcare Systems Engineering, vol. 1, pp. 20–36, 2011.
[27] M. Samorani and L. R. LaGanga, "Outpatient appointment scheduling given individual day-dependent no-show predictions," European Journal of Operational Research, vol. 240, pp. 245–257, 2015.
[28] D. Carreras-García, D. Delgado-Gómez, E. Baca-García, and A. Artés-Rodriguez, "A probabilistic patient scheduling model with time variable slots," Computational and Mathematical Methods in Medicine, vol. 2020, 2020.
[29] D. Elreedy and A. F. Atiya, "A comprehensive analysis of synthetic minority oversampling technique (smote) for

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3349953

handling class imbalance," Information Sciences, vol. 505, pp. 32–64, 2019.

[30] S. Tangirala, "Evaluating the impact of gini index and information gain on classification using decision tree classifier algorithm," International Journal of Advanced Computer Science and Applications, vol. 11, no. 2, pp. 612–619, 2020.

[31] R. Sharma, S. Gupta, and P. Garg, "Model for predicting cardiac health using deep learning classifier," in 2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT), pp. 25–30, IEEE, 2022.

[32] M. Jena and S. Dehuri, "Decisiontree for classification and regression: A state-of-the art review," Informatica, vol. 44, 2020.

[33] D. M. E. Barbosa and A. H. Ayala, "El uso de las tics en las pymes exportadoras," dimensión empresarial, vol. 15, pp. 121–143, 2017.

[34] A. Zendehboudi, M. A. Baseer, and R. Saidur, "Application of support vector machine models for forecasting solar and wind energy resources: A review," Journal of cleaner production, vol. 199, pp. 272–285, 2018.

[35] S. Dreiseitl and L. Ohno-Machado, "Logistic regression and artificial neural network classification models: a methodology review," Journal of biomedical informatics, vol. 35, pp. 352–359, 2002.

[36] M. M. Saritas and A. Yasar, "Performance analysis of ann and naive bayes classification algorithm for data classification," International Journal of Intelligent Systems and Applications in Engineering, vol. 7, pp. 88–91, 2019.

[37] S. MN and R. RW, "Improving accuracy metric with precision and recall metrics for optimizing stochastic classifier," in 3rd International Conference on Computing and Informatics, 2011.

[38] J. Huang and C. X. Ling, "Using auc and accuracy in evaluating learning algorithms," IEEE Transactions on knowledge and Data Engineering, vol. 17, pp. 299–310, 2005.

CATALINA VALENZUELA-NÚÑEZ received the M.S. degree in Industrial Engineering from University of Bío-Bío, Chile, in 2021 and the bachelor's degree in Civil Industrial Engineering from University of Bío-Bío, Chile, in 2021. Since 2022, she has been a researcher, assistant professor and project engineer at the University of Bío-Bío. In July 2023, she was a speaker at the 23rd Conference of the International Federation of Operational Research of Societies in Santiago, Chile, in the "Knowledge, Technology, and Innovation" section. Her research areas are linked to machine learning techniques and optimization.

GUILLERMO LATORRE-NÚÑEZ received the M.S degrees in Engineering Sciences in 2018 from Pontificia Universidad Católica de Chile and Industrial Engineering in 2012 from University of Bío-Bío. Also the PhD. degree in Engineering Sciences in 2018 from Pontificia Universidad Católica de Chile. He is currently subdirector of the subdirectorate of postgraduate studies at the University of Bío-Bío. His research spans a wide range of optimization areas delves into addressing complex scheduling challenges in healthcare environments, such as operating room scheduling that takes into account all resources, post-anesthesia beds, and emergency surgeries. Furthermore, he has explored efficient scheduling methods for hazardous materials collection with multiple-product loading and devised a heuristic approach to optimize the scheduling of mobile dental clinics while emphasizing fairness among school districts.

FREDY TRONCOSO ESPINOSA received the PhD. degree in Engineering Systems in 2015 from Universidad de Chile, Santiago. The bechelor's degree in Civil Industrial Engineering from University of Bío-Bío, Chile, in 2002. He is currently director of the institutional analysis department of the University of Bío-Bío. His research encompasses diverse areas of investigation, with a particular focus on utilizing machine learning techniques. Notably, he has developed a predictive model for customer churn in a telecommunications company in Chile, aiming to enhance customer retention strategies. He has also explored the potential of telemedicine for healthcare administrators, contributing to revolutionizing the healthcare industry. Moreover, his work has concentrated on identifying key individuals and their interactions within criminal clusters, utilizing attribute-based analysis within a social network context.