Predictive Scheduling and Throughput Optimisation in Outpatient Clinical Workflows: Machine Learning Models for Ambulatory Care Efficiency

Dr. Ana Castaño, Associate Professor of Computer Science, University of Buenos Aires, Argentina

1. Introduction

Machine Learning Approaches for Improving Patient Flow and Efficiency in Ambulatory Care: AI Models for Streamlining Scheduling and Reducing Wait Times

The value and importance of healthcare services are realized through care delivery; therefore, it is vital to seek innovations for increasing operational efficiency and patient access to care. Artificial intelligence (AI) provides a way to achieve this endeavor. For example, AI can improve capacity management through predictive modeling, as well as scheduling and patient flow in ambulatory care. In busy ambulatory clinics, long wait times encourage no-shows and lower patient satisfaction, while scheduling backlogs hinder access to new patients. Many clinics can address these challenges by increasing appointment capacity or flattening demand using customized nurse or physician schedules interspersed with higher-capacity templates or time cushions. Others invariably flip between overcapacity and undercapacity because their skill mix, appointment duration, and consultation content do not allow for custom scheduling; we focus on these clinics.

Machine learning (ML) has not been widely explored in ambulatory care scheduling research or practices. Consequently, the status quo of scheduling is largely population-based. In such systems, patients are scheduled at a standardized frequency based on the acceptance rate of appointment requests. Notably, twice as many clinicians use limited daily appointment availability for each type or urgency of appointment requests, rejecting the rest, as compared to making appointments for all request arrivals at a later date. Unfortunately, both systems disadvantage patients. Rejection prioritizes certain patients or slots based on appointment type or order of request; providers justify rejecting appointment requests to match schedule supply with providers via artificial demand restrictions. On the other hand, once a healthcare provider decides to book an appointment, the patient's pathology, how long they have waited for the appointment, their reason for visit, and the specific provider requested can be reduced to a scheduling-to-obtain ratio. The decisions made by a scheduling-to-obtain ratio are disparate from personalized or equitable care, and the scheduling-to-obtain optimization of the schedule varies widely based on an institutional culture of wait, among other factors.

1.1. Background and Significance

Patient flow and efficiency in ambulatory care

Several factors contribute to the need to ensure patient flow in ambulatory (outpatient) care is efficient and effective. The number of ambulatory visits is at an all-time high and it will only continue to increase. Many of these patients seek care to prevent needing more expensive urgent or emergency care treatments. Research has shown that patients who attended accident and emergency, compared to those who instead visited a general practitioner or ambulatory care, were more likely to rate care as poor, trust staff less, and be less satisfied with the information they were given. On the other hand, a positive ambulatory care experience increases patients' perception of quality of care and trust in their health systems. Despite this knowledge, in September 2021, almost 1.5 million people in the United States visited urgent care due to unavailability of appointments or long wait times in ambulatory care.

However, ambulatory care resources—including physical space, operating room time, and staff—are constrained and budgets are tight, creating a bottleneck that regrettably impacts patient experience. A critical part of improving patient flow and enhancing patient care experiences hinges on scheduling patients with accuracy. Schedule accuracy or appropriateness indeed has a direct link to lower waiting times and appointment no-show rates. By extension, using predictive modeling or AI to create appropriate care appointments can increase accessibility to care and improve service and care quality. Despite the differences between physician practices and health systems, all entities must schedule to meet the intrinsic demand that is shaped by variations in both population incidence of conditions and patient severity at any given point in time. In this variance, however, also resides an opportunity to stagger demand and allow patients to wait without visibly waiting. This will be the focus of the methods developed and the general framework for implementation.

2. Challenges in Ambulatory Care

Patient flow in ambulatory care poses many problems. Developing optimal templates for coordinating patient assessments is challenging because of high-dimensional appointment transitions. A suitable inventory of providers and resources to coordinate services for a wide range of patients at different levels of acuity is required. Appointments require scheduling for resources, being with a specific provider if necessary. Additionally, the range of services offered may vary significantly based on location. Patients' needs and optimal scheduling strategies are not the same for all patient groups, but many appointment systems cannot be easily modified to reflect those differences. Inefficiencies in patient coordination can also be caused by other forms of "waste," defined in healthcare as activity that consumes resources without adding value. This could include excessive paperwork, needless phone calls, and "no-show" patient appointments.

The types of service that can be processed within the appointment structure are limited by available resources. The strings associated with complex flows to provide these services may require different amounts of provider time. This, coupled with a diverse patient population with competing and complementary demands, results in longer waiting times and potential delays in care delivery. To overcome this, scheduling inefficiencies are common, and in some cases are enforced by regulatory agencies. One approach used in healthcare to manage uncertainty and increase efficiency is overbooking. This causes patient appointment times to widen, accompanied by an increase in work for many but not all providers. Overbooking, though used, is generally not ideal because of the reduced flexibility and additional costs. Inefficiencies in the patient coordination process can also be detrimental to the patient, with unacceptably long wait times being a contributing factor to less patient satisfaction. There are financial consequences for inefficiencies as well, with scheduling delays and loose schedules causing revenue deficits for institutions.

2.1. Patient Flow and Efficiency Issues

Fast, efficient patient flow is integral to the timely delivery of ambulatory care. The objective of patient flow is to ensure a patient progresses through examinations and treatment without any impediments, arriving at the appropriate hospital department with the necessary staff and medical equipment already in place and knowing what procedure they are to have performed. A good patient flow management system conveys accurate information with good timing, appropriate technology, and excellent customer service, promoting patient satisfaction and trust in care—two areas of great importance for providers, patients, and hospital management. Many factors can hinder patient flow efficiency. Prolonged waiting is the source of much patient dissatisfaction. While patients tend to expect long waits in emergency departments, an unacceptable wait is considered to be about 40 minutes. For patients in operating room recovery, a post-operative interval of 1.3 hours is considered to be an unreasonable wait time. Average patient wait times in ambulatory services are reported to have important implications for practice. The scheduling of just one additional patient per session was linked to reductions in patient wait times and the proportion of patients who left the practice without being seen.

Other inefficiencies associated with ambulatory care delivery involve the clinic administration. Scheduling is believed to be an inherently complex problem that can be influenced by a variety of resources, policy, and environmental constraints. Optimal clinic scheduling has the potential to support improved patient flow, clinic efficiency, and provider satisfaction while still accounting for patient needs and preferences. A lack of early morning appointments may lead to time differences between arriving for and leaving from hematologic or oncologic clinics, which can detrimentally impact work productivity and overall patient satisfaction. Late appointments are perceived as low in demand and poorly organized. Many patients pre-appointed for the final time slot are accustomed to excessively long delays while at the clinic. Persistent waiting and scheduling issues are stressful for patients. Gynecologic oncology patients have called for an earlier appointment time or lead time for computed tomography scans of the abdomen and pelvis. They expressed that it is maddening to schedule your entire day around an afternoon/evening appointment and then not get in until several hours later. Following an oncology multiphasic screening exam, one patient despised the inefficient logistics in healthcare, waiting to find out your results is the worst. I made my appointment when I was there. I was in that screening all day. I could have fallen asleep waiting for my result. Every day after leaving the clinic without the results of my MRI, I was thinking two words: wasted time. As an inpatient or day care cancer patient, I feel like time I should be able to get back to my life is lost.

3. Machine Learning Applications in Healthcare

With the growing surge of data, particularly in healthcare, hospitals worldwide are scrambling to utilize and deploy machine learning and artificial intelligence (AI) to augment health service delivery. Machine learning has the potential to transform the healthcare industry by providing predictive analytics, risk stratification, and high-quality decision support. The employed models can be vast and vary from classical statistical models, decision trees, and ensemble models to more recent deep learning methods. AI presents a key area for investment and use in the resolution of major healthcare delivery problems.

Machine learning and predictive analytics provide a substantial foundation for efficient patient-centered healthcare. Although recent developments in AI applications have projected significant gains when compared to their human counterparts, one must consider the ethical considerations and limitations behind them before adopting them into patient care. These considerations might involve data privacy concerns, algorithmic transparency, and interpretability. AI can be largely divided into three categories: supervised techniques requiring large labeled datasets, unsupervised techniques without labeled datasets, and reinforcement techniques, combining both. AI and machine learning have multidisciplinary interactions with other technologies applied in healthcare, such as the Internet of Things, cloud computing, edge computing, or fog computing. In particular, these technologies influence the decisions for model deployment, training, and development.

3.1. Overview of AI and Machine Learning

Given the widespread enthusiasm and interest in AI, there is unavoidable confusion about what AI consists of, and specifically, what is meant by "machine learning." AI refers to systems or machines carrying out tasks that would typically require human intelligence to sort through, interpret, and assess. AI involves problem-solving, planning, perception, recognizing speech, understanding natural language, and so much more. This interdisciplinary field leverages neural networks, information engineering, computer science, psychology, and many additional disciplines. Machine learning can be generally defined as the extraction of patterns and information from enormous data. More specifically, machine learning consists of the method in which an AI example utilizes sets of data to uncover various kinds of associations or insights. It is critical to remember that all machine learning is, in fact, AI, but not all AI consists of machine learning. Machine learning is modular in the sense that some systems depend on extensive data sets. Furthermore, there are three central learning processes: supervised learning, unsupervised learning, and reinforcement learning.

In supervised learning, AI models are taught what output they ought to produce for a given data input. In unsupervised learning, AI models are trained to interpret data themselves. If we applied unsupervised learning to the hourly traffic data in a heavily congested metropolis, the clusters of traffic patterns may arise. Lastly, in reinforcement learning, AI models are given a task to complete independently. Medical professionals can use AI in diverse settings, but in general, the goal is to gain insights and speed up efficiency. A variety of algorithms can be employed in conjunction with AI, depending on the desired results. Some collect data and make predictions, some uncover relationships among data, and some employ external reinforcements to make decisions. AI is, in essence, utilized to sift through large data sets to extract and make use of the intelligent insights buried within. Large data is a major issue in ambulatory care, and AI is expected to play a critical role in this sector in the upcoming years.

4. AI Models for Scheduling Optimization

The heart of scheduling is the idea of optimally allocating resources, and successful scheduling reduces queues and lessens wait times. Appropriate staffing of doctors and nurses, prompt assignment of diagnosis and treatment, and the rapid turnover of patients within clinics are fundamental to efficient scheduling. Increasing research interest in improving patient flow and efficiency confirms there is a promise in developing accurate patient flow and efficiency estimators, since these will be used for system design and simulation, demand management, and more efficient rostering of medical staff. In the adjacent section, we explore AI models of outpatient appointments scheduling in more detail.

While scheduling was traditionally performed using simplistic approximations, current research now encompasses data-driven models for both prediction and optimization.

The prediction approach estimates the number of patients who might arrive for care at different times, and this information is subsequently employed to prioritize incoming patients and estimate future workload. Predictive models use time series regression. In addition to arrival rates, predictive modeling can also be used to predict other factors, such as patient demand, care duration, visit complexity, and patient waiting times. On the other hand, a scheduling policy will determine how incoming patients will be seen by aggregating desirable characteristics into functions that reflect providers' performance.

These scheduling solutions usually seek to fill the day's staff and resources as evenly as possible so that offices run at or near full capacity, to create a fixed allocation of visits to providers, or to estimate schedulers' preferences for priorities in real time. Unlike existing solutions, AI models offer a number of advantages. They consider not only priority but also estimate willingness to reschedule and likelihood to no-show and thus balance different patient behavior and clinic performance. The results are real-time decision models that provide scheduling through prioritizing, redistributing, or construction of a slack schedule. Machine learning models have succeeded in anticipating point-of-care interruptions and predicting no-shows and cancellations. Moreover, machine learning models can learn and model patient behavior, and provide preference-justified and individualized scheduling solutions for different patients, for example in clinical trials, complex conditions, or disease onset prediction. In the following sections, we will provide a brief review of such AI models that are developed or applicable in ambulatory care.

4.1. Types of Scheduling Algorithms

Rule-based algorithms are most prevalent in practice and are based on a sequence of logical rules developed by an expert or by observation from operational data. Although they are entirely transparent, adaptive, and unobtrusive, they can become complex to change or expand due to an order dependency among rules, and a single global performance criterion is usually ignored. Nonetheless, they are adaptive and can accommodate case-specific inputs, leading to an increase in their use; for example, one case study showed how such algorithms can be used to differentiate between case instances characterized by different patient urgency levels and provide priority to the most urgent cases in this regard. Optimization-based scheduling algorithms use a performance measure as an objective function that needs to be minimized or maximized. Although such algorithms are generally good at managing repetitive flow processes, such as those arising in manufacturing systems, or for optimizing specific allocation services such as workforce planning given a fixed demand, it can be more difficult to use an optimization-based approach to manage patient care.

Alternatively, machine learning-based scheduling can be used in preference to rule or optimization-based strategies when a combination of data sources is available to adjust scheduling decisions according to evolving patient flows. A strength of machine learning-based algorithms is their potential to model nonlinear phenomena or interactions captured from the raw or processed input signal. On the other hand, the algorithm is a black-box system affected by unobserved dynamics or dependencies. As a result, for healthcare providers to reconsider and potentially replace their well-functioning rule-based scheduling systems with a learning-based alternative, a different value proposition needs to be offered. Learning-based algorithms can be directly implemented into the scheduling solution, leading to the corresponding output of the scheduling policy, instead of suggesting changes to the input signal. This can lead to more time-efficient scheduling and a reduction in the manual adjustment of an already functional decision-making system. Several studies reported successful applications of machine learning-based scheduling algorithms for healthcare operations. In outpatient services especially, demand is highly variable, patient behavior is increasingly difficult to model, and patient wait times are of high importance. For these reasons, this paper largely focuses on machine learning applications in and design methods for outpatient clinics.

5. AI Models for Reducing Patient Wait Times

Wait times are an important determinant of patient satisfaction and quality of care in healthcare systems. As a result, there has been significant interest in AI models to address scheduling and patient flow to improve wait times. AI has utilized data-driven decision-making and analytics to develop strategies to improve scheduling, further reducing wait times. One important area of research for many AI studies is the anticipation of patient flows to prepare for surges in patient arrivals. AI systems look to plan for short-term surges in arrivals for maximal operating ability in procedures and emergency treatments.

A key application of AI in outpatient settings is to streamline appointment scheduling and anticipate patient no-shows or late arrivals. Active research has been conducted to use AI to create appointment templates. Using data analytics, AI created appointment schedules that accommodate more patients with templates and the ability to see additional patients. Hospitals that have implemented data-driven scheduling algorithms have seen success in reducing wait times and better utilizing surgical block time. As waiting room congestion can have important spillover effects on resource utilization, other AI models improve patient flow or help solve downstream challenges. For example, AI has been used to initially improve patient flow in an emergency setting with success in reducing wait times, ambulance diversion, and left without being seen rates.

One of the most significant challenges in regard to utilizing AI to anticipate, influence, and adapt to changes is the time scale of operations in an outpatient setting. Not only is the time scale of operations on a much finer resolution, but the schedules and workflow change day-to-day in a method that is less computationally intensive to model. In outpatient clinics, AI is used to forecast patient arrivals and adjust staffing needs in light of the forecasted patient flow. In one instance, the outpatient clinic utilized AI to deploy forecasted staffing needs, resulting in reductions in patient wait time, length of stay, and increased patient satisfaction in both wait time and clinic flow.

5.1. Strategies for Wait Time Reduction

In an effort to reduce patient wait times in ambulatory care settings, numerous strategies have been explored. Machine learning has been used to determine peak times, improve scheduling, and increase efficiency in the system. Most commonly, machine learning techniques use historical no-show data for predictive analytics, rerouting, scheduling optimization, and flow management. Wave analysis techniques have also been developed. However, these techniques target a static problem and do not adapt based on new information. Certain flow management strategies have been used in healthcare approaches.

Because wait time is commonly defined as patient flow, most literature focuses on flow management for efficiency and targets operational changes, but there is minimal literature on administrative changes and improved access in patient care. Overbooking involves making additional appointments than can be physically handled in order to reduce lost revenues due to no-shows or last-minute appointment cancellations. Tiered scheduling coordinates the appointment calendar where slots used for appointments are tiered for different needs, and the remaining slots are reserved for higher-acuity patients, rendering emergency departments not needing appointments. Continuous feedback loops using real-time data ensure that model predictions or allocations are still valid and act as a metric for scheduling accuracy.

In ambulatory care, these interventions have resulted in substantial wait time reductions. Implementations of learning health systems in healthcare organizations may present challenges in the form of hesitance in the workload of system staff, technological ability, and administrative leadership. By illustrating successful implementations and reductions in ambulatory care, these strategies support technology-enabled approaches in ambulatory care to improve access and efficiency. Time-series algorithms are able to consider a wide range of data and analyze points of wastefulness in the system. Random forests are non-parametric models that can capture non-linear behavior with independent parameters. This makes it the ideal model for learning from priority scores in the random forest: stage two users over different months and understanding what defines high access times.

6. Future Direction

The unprecedented advancement in data science and artificial intelligence (AI) has fueled the revolution of predictive analytics. In combination with real-time technologies such as IoT, data is being generated and processed faster than ever. This section will now present emerging trends and the future direction in both healthcare management of ambulatory care from a patient flow and efficiency perspective, as well as enhanced data-driven scheduling. In the section of healthcare management and patient scheduling, we will then present potential future goals, methodologies, and technological advances that could assist in their achievement.

The influence of AI-based techniques will have a profound impact on the current healthcare delivery systems. This trend is powered by the rapid advances in big data and AI and has been creating innovations in the electronic health record (EHR) as well. While predictive algorithms based on regression and decision trees have seen some adoption in strategic planning in healthcare systems and medical management of patients, it has also been argued that potential benefits of more advanced AI for healthcare delivery are yet to be realized. It is known that these scheduling methodologies can reduce patient wait times while occluding capacity, thus tarnishing performance outcomes of healthcare systems. Indeed, intensive study argues that too aggressive a scheduling system can be harmful to the delivery of healthcare systems. However, with a move towards a data-driven and personalized healthcare system, it will be crucial for the health administrative staff to consider the needs of both the system in need of delivering the care as well as the patient who receives it. This entire line of work is at an early stage and so requires much future research and development. We expect future research to encompass a wide array from theoretical new methodologies in AI to fully fledged health systems research and evaluation. Furthermore, the transitioning of a digitally augmented, data-driven, and AI-compatible healthcare system will face challenges. These can be as mundane as data sharing and privacy from separate health systems to increasing concerns about the ethical use of AI and the potential corruption of AI output. Thus, as specialist working groups that include stakeholders such as researchers, care providers, hospital administrators, and policymakers have been set up in the scheduling community, we propose that this should soon follow suit in the wider healthcare machine learning community. As algorithm innovation blossoms, and a foundation of top-down engagement is fostered, we expect research in healthcare machine learning to foster future health directions.

7. Conclusion

The advances made in machine learning and AI can be harnessed to make scheduling appointments more efficient and less time-consuming for both staff and patients. Furthermore, once we bring patients into the clinic, AI models can help improve patient flow and reduce wait times. There are various challenges that manifest in the form of long waits—peaks in appointment request calls and high variability in appointment duration are two such examples. Through machine learning algorithms, we can work towards improving the way appointments are scheduled and smooth out some of the inherent inefficiencies associated with health care. Existing research suggests reductions of up to 45% in patient time in the system with a corresponding reduction in no-show rates of up to 70% as a result of the integration of AI-based scheduling.

We express the importance of AI system development from the perspective of patient scheduling and complex ambulatory care, where much of the healthcare business is being—and increasingly will be—conducted. Research is intended to benefit the most significant areas of patient care, and seamless scheduling and decreasing time spent waiting for care are identified as essential parts of increasing the quality offered to patients. Earlier research has revealed and justified the application of various AI models to assist in this area through predictions about pre-appointment length, appointment type, and no-shows. It is continuously imperative to research novel approaches, as many of the outdated research systems are suitable only in their own right and are not very suitable for healthcare systems today. Future artificial intelligence research should focus on collaboration with cross-disciplinary groups toward systems to schedule all forms of patient care, including yet unaddressed periods between appointments.

In conclusion, machine learning optimization models which leverage advanced predictive models can help arrange patient systems and reduce patient time spent at healthcare clinics on the day of the appointment. Clinically, we believe that our model represents an achievable approach for managing health clinic systems. It has the potential to lead to important benefits for both patients and clinicians in healthcare settings. Of additional interest would be to explore how reducing waiting times and improving the clinic experience impacts other aspects of a health clinic that have been studied previously. While changing appointment scheduling and patient flows based on these structured results would require more elaborate testing, this study offers a technologically driven policy pathway for many clinics today. Ultimately, many areas of the healthcare system have remained beyond the reach of AI systems. AI-based assistants have been developed for radiology findings and for aiding primary care doctors in making complex diagnoses. Going beyond staff-side scheduling tools, the area of ancillary patient services scheduling and scheduling systems aimed at patient-facing scheduling systems generally have received little dedicated attention. We believe that future research should focus on interdisciplinary approaches to improving patient scheduling and patient care. The underlying approach to using predictive data drawn from massive data repositories to improve day-of patient scheduling appears to be both immediately accessible and to offer quick return.