Vol. 07, No. 3 (2025) 1601-1614, doi: 10.24874/PES07.03.021

Proceedings on Engineering Sciences

www.pesjournal.net

IDENTIFYING INFLUENTIAL FACTORS BEHIND PHYSICIAN PERFORMANCE: A MACHINE LEARNING APPROACH TO SUPPORT DECISION-MAKING

Amani Mustafa Ghazzawi[1]
Alaa Omran Almagrabi
Hanaa Mohammed Namankani

Received 19.11.2024.
Revised 07.04.2025.
Accepted 24.04.2025.

Keywords:

Physician, Data mining, Performance, Medical, Evaluation, Error, Decision, Linear Regression

A B S T R A C T

Performance improvement is essential to every organization including medicine. For Human Resource to evaluate levels of performance quality, standards should be established as a base of developing evaluation methods. This paper aims to facilitate and enable more efficient performance evaluation and improvement methods for physicians by discovering the most reliable features. It will identify the main features influencing physicians' performance including line managers, patients, and physician information. These features are represented in a use case design that illustrates performance appraisal behavior. The paper presents a data set that includes 500 physicians in an Egyptian hospital, which was created based on the proposed features, and features affecting the physician's performance will be tested using Regression analysis. Both the Lasso and Ridge regression models showed significant performance improvements. Lasso achieved an R^2 of 0.413 and an RMSE of around 0.01264, while Ridge achieved a slightly lower R^2 of 0.397 and an RMSE of around 0. 01281. The improved performance of these models can be attributed to their incorporation of regularization terms. Finaly, if these features are extracted automatically, they will better support the decision-making process. These features may help design a framework that helps evaluate physician performance and supports decision-makers in the future.

1. INTRODUCTION

Aiming to evaluate physicians' performance as professionals, several healthcare systems are responding to the need for high-quality treatment. The requirement for external assessments arises from the restricted capacity for self-assessment. To support these efforts, it is essential to have performance metrics that are valid, realistic, dependable, and efficient. Since 1993, health systems all around the globe have been using multisource feedback (MSF), also known as 360-degree assessment, more often to evaluate various aspects of professional performance. MSF entails the external evaluation of physicians' performance on a range of duties by three

[1] Corresponding author: Amani Mustafa Ghazzawi
Email: amani.ghazzawi@yahoo.com

parties: patients, non-physicians (nurses, allied healthcare professionals, or administrative personnel), and peers with expertise in a comparable area of practice. To be able to respond to inquiries on a physician's efficiency, raters in those three categories had to watch the behavior of the medical professional. To determine areas for improvement, physicians are also asked to rate their performance on a questionnaire. These evaluations are then compared to those of others. Physicians, management, and patients must have faith in the validity and reliability of the instruments used in MSF, as this trust is the cornerstone of the process. The validity and reliability of the instruments used for MSF have been demonstrated in both the UK and Canada across a range of specializations. Nevertheless, a recent Campbell et al. (2008) study discovered that significant sources of systematic bias, such as a physician's specialization and whether they serve in a locum role, impact these multisource evaluations. This suggests that a physician's MSF score may be influenced more by the responder's socioeconomic background than by the physician's actual performance, which must be examined in a variety of MSF contexts.

The medical profession contributes to the building and development of society. It is one of the professions that most need awareness of conscience because the physician is entrusted with the lives, secrets, and symptoms of people. The outbreak of "Covid-19" asserts the importance of enabling the health sector to face any potential crises in the future successfully. From this point, this study highlights the most valuable element of the health sector: physicians. As Bindels et al. (2021) pointed out that the health sector has not received the required attention in some countries that are severely affected by widespread epidemics such as coronavirus. The study elaborates on the critical role physicians play in serving their societies and their efforts during these hard times challenging the world. All countries and communities appreciate these efforts. As Sargeant (2019) adds, physicians represent the first defiance line in facing global epidemics. In contrast, the most hospitals lack reliable methods to assess physicians' performance. Therefore, HR in medical organizations needs to find a methodology to effectively evaluate physicians' professional performance and detect their weaknesses for the continuance of performance improvement. Therefore, it was necessary to discover the features that affect the evaluation of the physician's performance. In this regard, this research tackles the role of "data mining techniques" in hospitals as an essential tool to support evaluating and improving physicians' performance and the HR decision-making process, which involves identifying areas for improvement, setting performance standards, and making decisions about promotions and incentives.

Similarly, this research work aims to detect the features that influence the process of evaluating and improving physicians' performance to efficiently enable decision-makers to make the right decisions and participate in reducing medical errors by allowing a more efficient evaluation process. Brennan et al. (2014), found that most studies in the health sector focused on evaluating the physician's performance manually through direct management, including professional and personal performance. The most prominent features used by Ho and Baker (2012) studies using the classical method are the physician's personal information (age - specialty - gender), medical knowledge, communication skills, peer evaluation, patient satisfaction, practical experience, and professional competence. Kuemmerle (2014) proved that these classical methods, though labor-intensive, have proven to be thorough and effective. The tools that was based on (Zhang et al, 2004) were regression analysis, integrating Norman's action theory and Reason's theory of human error, developing a framework for medical practice, a MOC program, and a systematic search of electronic databases, a system based on the principles of appreciative inquiry, Pearson's correlation coefficient and linear mixed models. However, there was a study that used AI to evaluate physician performance. Shi et al. (2018) used text-based online consultations between physicians and patients using Python programming with an ordinal partitioning scheme (SVMOP), and its features were based on the number of medical terms used by the physician, the number of questions asked by the patient, and predictive features, such as politeness and emotional words.

Moreover, many studies have used the computerized way of evaluating the physician's competence, not performance. There was a group of studies that used machine learning techniques of all kinds but were in assessing the physician's competence. Dias et al. (2017) relied on many features in the evaluation, such as patient care, medical knowledge, practice-based learning and improvement, interpersonal and communication skills, professional competence, and systems-based practice. These studies used many machine learning tools, such as Natural language processing, support vector machines, neural networks, and hidden Markov models. This study proposes a set of essential features that influence the evaluation of physicians' performance by analyzing their data (features) using data mining techniques and directing their potential through procedures that help support decisions to improve performance and reduce medical errors. The paper also presents the dataset that was taken from Egypt Hospital to apply the proposed features, demonstrating the thoroughness of our research.

1.1 Significance

Identifying influential factors behind physician performance is crucial for enhancing healthcare quality and patient outcomes. Factors such as individual, psychological, and organizational aspects play significant roles in determining physician performance. Additionally, factors like role balance, autonomy, managerial support, teamwork, and relational capacity are essential for success among physician-scientists in

Academic Medical Centres. However, mentorship provides the necessary guidance and support to help these factors become more effective. Understanding how physicians impact each other's performance during shifts can lead to improved hospital operations, as peers influence each other's speed and quality, especially during high-volume shifts. Analysing these factors through bibliometric studies can reveal trends in healthcare worker performance, with a change towards psychological factors like burnout and anxiety, particularly highlighted during the COVID-19 pandemic. By identifying and addressing these influential factors, healthcare institutions can optimize physician performance, ultimately benefiting both healthcare providers and patients.

1.2 Motivation

Most previous studies relied on manual assessment of the physician’s performance, underscoring the potential for automation and the need for more advanced techniques. Therefore, the researchers conducted a study that used artificial intelligence (AI) techniques to evaluate the physician’s performance. As it plays a crucial role in determining physician performance, with intrinsic and extrinsic factors influencing their behaviour and outcomes. Studies have shown that factors like recognition, achievement, nature of work, and job security significantly impact healthcare professionals' motivation and work engagement. However, it is not enough to just identify these factors. The urgent need is to understand and address these motivational factors, as they are essential for enhancing physician performance and ensuring the delivery of high-quality healthcare services. Additionally, the measurement of physician performance and the use of incentives have been critical focuses in healthcare, with performance measurement systems incorporating both extrinsic and intrinsic motivators to drive improvements in care quality. Furthermore, research indicates that intrinsic motivation, such as feelings of accomplishment and autonomy, can have a more decisive influence on employee performance compared to extrinsic factors like salary and working conditions.

Identifying influential factors behind physician performance can be challenging due to various reasons highlighted in the research. Factors such as individual, psychological, and organizational aspects play significant roles in determining physician performance. The complexity of assessing individual physicians within teams and systems of care is a substantial aspect of this research, and the uncertainty surrounding the role of financial incentives in motivating physicians adds to the difficulty of pinpointing influential factors. Moreover, the influence of context factors, including Physician-related, patient-related, and consultation-related aspects, can impact communication performance levels in healthcare professionals, making it harder to identify key determinants of physician performance. Additionally, factors like role balance, autonomy, organizational support, teamwork, and relational capacity have been identified as crucial for the success of physician-scientists, further emphasizing the multifaceted nature of physician performance assessment. Therefore, it has been revealed from the previous studies that the following limitations still exist:

- Most studies concerned with physician evaluation have used a manual method.
- It is crucial to recognize that a more effective evaluation method could significantly enhance the physician's performance, which is primarily assessed through manual peer evaluation, questionnaires, and feedback.
- Previous work mainly relies on interviews conducted in some hospitals, which is a manual evaluation method.
- There are no studies on evaluating physicians and predicting their performance, as most studies focus on assessing physicians' competencies, which are the skills and knowledge they possess. In contrast, assessing performance involves measuring how effectively these competencies are applied in practice.
- By identifying the best techniques for evaluating a physician's performance, we can empower decision-makers to make appropriate decisions that have the potential to reduce medical errors in the future significantly.

1.3 Main Contributions

The significant contributions of this work are as follows:

1. A specific data mining (DM) technique, such as predictive modelling or clustering, enhances the evaluation criteria and revolutionizes physician evaluation.
2. Researchers have used specific computerized methods, such as machine learning algorithms or data analytics tools, to evaluate physicians' performance.
3. Data mining reveals the factors influencing physicians' performance. This information can empower HR to make responsible decisions that could reduce medical errors.

2. RELATED WORK

This section presents existing works and methods indicated by researchers in evaluating the professional performance of physicians. It aims to reveal the features that influence the evaluation performance of physicians, and the techniques used in the evaluation. (Campbell et al, 2008) adopted a cross-sectional questionnaire survey of physicians, patients, and colleagues to collect information on physicians’ professional performance in the UK. The authors used several statistical methods to analyse the data, including principal component analysis, factor analysis, and regression analysis. The questionnaires assessed patient satisfaction with the physician's communication skills, clinical competence, and professionalism, as well as the colleague assessment of the physician's communication skills, clinical competence, and professionalism, and the physician's age, gender, and specialty. The researchers indicate that questionnaires are not an absolute independent measurement of performance. Overeem et al. (2012) also an observational validation study of three multisource feedback (MSF) instruments was also conducted to evaluate physician performance by peers, co-workers, and patients. In addition, physicians filled out a self-evaluation form. The authors used principal component analysis and classical test theory methods to evaluate the instruments' factor structure, reliability, and validity. The features assessed in the study are i) Peer ratings of the physician's communication skills, clinical competence, professionalism, and overall performance; ii) Co-worker ratings of the physician's communication, teamwork, and overall performance; and iii) Patient ratings of the physician's communication skills, empathy, and overall performance, as well as self-ratings of the physician's communication skills, clinical competence, professionalism, and overall performance. The study findings reveal that there is no correlation between self-assessment scores and the scores of colleagues, workmates, and patients.

As Bindels et al. (2021) conducted a system for performance appraisal for physicians in a Dutch Academic Medical Center. The system consisted of a two-hour peer-to-peer conversation based on the principles of appreciative inquiry and solution-focused coaching. The system included the following features: i) peer-to-peer approach, ii) professional development, iii) feedback, and iv) continuous improvement. Their work is still awaiting verification of its impact on professional development and presents findings that are particularly relevant to physicians. Sargeant (2019) indicated the use of feedback and coaching conversations in the context of professional development, a topic of great interest to our audience. The study identified the features of feedback and coaching conversations that are important for supporting the development of self-directed learners and resilient professionals, such as timeliness, where feedback should be provided as soon as possible after the event or performance that is being evaluated, the Specificity of Feedback should be specific and focus on the individual's behaviour or performance, rather than on their personality or character, also, the actionable Feedback should be actionable, the supportive Feedback should be supportive and respectful, even when it is negative. The coach should ask questions and facilitate reflection to help the individual learn from the feedback and develop their solutions. It is a survey article, so it does not include any research data.

Brennan et al. (2014), conducted in the United Kingdom demonstrated that the evaluation of physicians is essential to maintaining a professional license, improving clinical performance, and changing physicians' behaviour. The study uses a systematic approach to search for and select evidence. This includes searching electronic databases for published research and 'hand-searching relevant journals, a process where researchers manually review the contents of specific journals to identify relevant articles. The evidence will be analysed using realist logic to determine the mechanisms through which appraisal produces its effects. The study is based on a realist review methodology, which is a complex and challenging approach to evidence synthesis. The study will provide valuable insights into how the appraisal of physicians produces its effects.

In a study conducted by Ho and Baker (2012) the General Medical Council (GMC) set a framework for good medical practice by integrating performance evaluation for the re-certification of physicians, which includes verifying the validity of licensed physicians every five years and demonstrating their ability to practice. The GMC framework for appraisal and revalidation considered i) knowledge and skills, ii) communication and interpersonal skills, iii) decision-making, iv) patient-centered practice, v) teamwork and collaboration, vi) management and organization, and vii) professionalism and ethics. However, their research is documentary proof and does not provide any empirical evidence. Kuemmerle (2014) pointed out that the American Board of Medical Specialties (BMS) enriches the Maintenance of Certification (MOC) program with standards that are pivotal for the professional development of medical practitioners. This program, which is a cornerstone of their professional roles, significantly contributes to the improvement of medical practice and patient care through continuing medical education (CME), performance improvement, and assessment. The MOC program meticulously evaluates the following aspects of physicians: Medical knowledge, Patient care skills, Communication, interpersonal skills, and Professionalism. This comprehensive assessment inspires and motivates medical professionals to strive for excellence in their practice. Zhang et al. (2004) proposed a cognitive model that categorizes medical errors based on the level of the individual and the extent of his interaction with technology. The study developed the cognitive taxonomy of medical errors by tool integrating Norman's action theory and Reason's theory of human error. Norman's action theory describes the steps involved in human goal-directed behaviour. Reason's theory of human error distinguishes between slips and mistakes. Slips are unintentional errors that occur during the execution of a planned action.

Moreover, Shi et al. (2018) introduced a unique approach, employing ordinal regression, to evaluate physician performance from online textual consultations. This strategy involves several key steps, starting with feature engineering. Here, we extract a diverse range of features from the textual consultations, including basic text features, customized features, and predictive features. The study utilized the Python programming language and the sci-kit-learn library to implement the model. An ordinal regression algorithm was employed to train the model, which was achieved by training a support vector machine combined with an ordinal partitioning scheme (SVMOP) on the extracted features. Basic text features encompass the number of words, the number of sentences, and the average word length. Customized features, such as the number of medical terms used by the physician and the number of questions asked by the patient, were also considered. Predictive features, such as politeness and sentiment words, were included in the model. Similarly, the result conducted using data from an online medical consultation platform offers valuable insights into evaluating physician performance. However, it's important to note that the results may need to be more generalizable to other settings, and the study's reliance on self-reported data may introduce bias. Their study also does not consider all the factors that may influence physician performance, such as patient characteristics and the healthcare system in which the physician works. This is a limitation that future studies should address. Dias et al. (2017) stated in his study that many studies have used artificial intelligence techniques to evaluate physician efficiency. A systematic review of the literature was conducted in 2019 that sought to identify studies that used machine learning to assess physician competency in 69 studies. The most prominent features that were used to evaluate physician competency using machine learning were patient care, medical knowledge, practice-based learning and improvement, interpersonal and communication skills, professional competence, and systems-based practice. The studies included a variety of machine learning algorithms, including decision trees, support vector machines, random forests, and neural networks. Their study indicated that the specialty studied the most was surgery—general and radiology. The study pointed in the impact of learning techniques and their ability to integrate and analyse information that can be used in the evaluation process.

Mirfat et al. (2018) argued that individual, psychological, and organizational factors are all crucial in understanding physician performance. They found that psychological factors have the most significant direct effect, while organizational factors show a positive but not significant impact. Cassel & Jain (2012) suggested that intrinsic motivators like accomplishment and patient appreciation, as well as extrinsic factors such as financial rewards and recognition, play a key role in motivation levels. Cola & Wang (2017) demonstrated that role balance, autonomy, organizational support, teamwork, mentorship, and relational capacity are all crucial for physician scientists' success. These multifaceted factors are essential for understanding and enhancing physician performance in Academic Medical Centers. Jin et al (2024) also pointed out that the factors influencing healthcare workers' performance, including burnout and anxiety, were analysed pre and post-COVID-19, providing insights into improving physician performance. William et al. (2022) identified 22 variables affecting medical lecturers' performance, with leadership, commitment, and credit points being the most influential factors, showcasing the complexity of interrelationships in performance evaluation.

Moreover, Edo et al. (2023) said that limited consultation time, unfavourable working environments, and physician communication skills are key factors influencing physician performance in healthcare service delivery, as revealed in the study. Eva (2018) said that the factors influencing physician performance are complex due to subjective observer perspectives, cognitive limitations like attention and memory, and the profound impact of prior experiences, making assessment challenging in healthcare settings.

Therefore, it has been revealed observed from the previous studies that the following limitations still exist: Most studies concerned with physician evaluation have used a manual evaluation method. However, a promising solution is on the horizon- using data mining technology to enhance the evaluation criteria. This innovative approach could revolutionize physician evaluation. The second limitation concerns assessing the physician's performance, which was based on most studies on manual peer evaluation, questionnaires, and feedback. The researcher found only one study that used the computerized method to evaluate the physician's performance. Third, the researcher conducted interviews in some hospitals. These interviews reveal that these hospitals rely on a manual evaluation method. The fourth limitation refers to a need for studies on evaluating the physician and predicting his performance, as most studies focused on assessing the physician's competence. The fifth limitation views difficulty in evaluating a physician's competence without assessing his performance, so using the best techniques in evaluating the physician's performance must be highlighted by discovering the characteristics that affect the evaluation of the physician's performance and helping the decision maker in making appropriate decisions, that can reduce medical errors in the future.

3. PROPOSED FEATURES FOR EVALUATION

After reviewing previous studies, evaluating the data of physicians in three hospitals in the Kingdom of Saudi Arabia, and conducting personal interviews, the study proposes a set of features that are practical and may be used to evaluate a physician's performance and help support the decision. This proposed research methodology is based on a use case in Figure 1.

The use case includes the features used in physician evaluation: Personal Information, Professional information, External evaluation, internal evaluation, medical knowledge, systems practices, research projects and presentations, and physician's performance. These features, when considered collectively, ensure a highly accurate assessment of the physician. This information supports the decision maker in making the appropriate decision regarding which physicians may be promoted, trained, or recommended to attend conferences and write scientifically, with confidence in the reliability of the process. The same task was previously accomplished through a manual evaluation along with the addition of other significant features from physician evaluation forms in Saudi hospitals that may affect performance evaluation. So, the researcher, while personally interviewing field experts, concluded that hospitals in Saudi Arabia are some features to be added in evaluating the Physicians, such as the fact that the Physician is assessed by the Ministry of Health or the Saudi Commission for Health Specialties, as it became clear to the researcher that there are two bodies for evaluating Physicians. The method of evaluation differs for each.

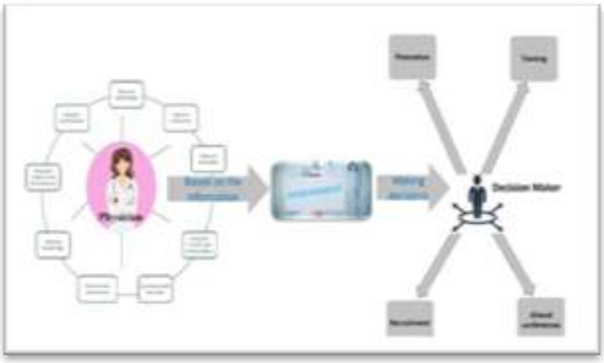

Figure 1. Use Case of Physicians Performance Evaluation

In fact, after studying medicine for six years at the university, the student then goes for a one-year internship, where they are evaluated by the consulting physician based on the evaluation criteria. The criteria were specified by the university, which sends the students to the consulting physician. The evaluation criteria of each university are different from those of other universities. Upon completion of the internship year, the medicine student has to choose to go to the Saudi Commission for Health Specialties (Saudi Board) or further for a subspecialty or continue as a general practitioner and is classified by the Ministry of Health as a civil service that covers any missing specialist in the hospital. The Ministry of Health is responsible for appointing and evaluating general practitioners and emergency physicians. The general physician covers any missing specialist in any medical field, and the emergency physician is assessed by the head of the department, which the Ministry of Health accredits.

Meanwhile, the Saudi Commission for Specialties appoints and evaluates consultant physicians, specialists, and resident physicians. The resident physician is appointed and assessed when the physician finishes the specialty of study and obtains a bachelor’s degree and distinction. If anyone wants to specialize in surgery, for example, they must join the Saudi Commission for Health Specialties and be tested, considering the rate calculation. They are then placed in the appropriate hospital. The Saudi Commission for Health Specialties appoints a consultant physician, a crucial figure in the process, to evaluate the resident physician and be responsible for them during the years of surgery for 5 years. The resident physician is trained with many consultants in many hospitals, where he is evaluated every three months by a specific consultant within a particular hospital. The resident physician does not have the right to perform any operation or open a clinic, so they must be under the supervision of a consultant. After five years, the resident physician receives the Saudi Board and becomes a specialist. The specialist practices the profession for two to three years, after which it is possible to become a consultant.

Table 1 briefly describes physician features that aid in evaluating their performance, along with data descriptions and the rationale for choosing each feature. These features, when adequately assessed and analysed, provide the decision-maker (HR) with the necessary tools to make appropriate decisions for each physician. This empowerment of the decision-maker (HR) is crucial in reducing medical errors. Features that have no effect are also deleted, further highlighting the integral role of the decision-maker (HR).

Table 1. Physician features in evaluating performance

Feature	Type	Description
Identity	Integer	A field with unique values in a table. Because each record has a different key value, key values can refer to entire records.
Age	Integer	(Campbell et al, 2008) adopted a cross-sectional questionnaire survey of physicians, patients, and colleagues to collect information on physicians' professional performance in the UK. One feature included the physician's age, gender, and specialty.
Gender	Character	
Nationality	Citizenship	We need this feature because the contracts of foreign physicians differ from those of national physicians in some countries or health institutions. These differences depend on local and national policies and regulations and the cultural and legal context of the country. Some foreign physicians need additional training under supervision.
Qualification	Character	A physician's academic and training qualifications are a significant factor in their evaluation. If a physician is qualified and competent in their field, they are more likely to receive a positive review.

Position title	Consultant physician Resident physician Specialist physician Physician intern General Physician	**(**Overeem et al, **2012)** working Experience and Rank were the variables that contributed most to performance increases.
Experience in Year(s)	Date/Time	To find out which departments are most exposed to medical errors
Department	List	
Saudi commission for health specialties	Yes \ No	Due to the difference in the evaluation mechanism between the Saudi Commission and the Ministry
Ministry of Health	Yes \ No	
Patient Assessment	{1,2,3,4,5} 1 = very poor 5 = very good	(Campbell et al,2008) adopted patients' and colleagues' questionnaires to collect information related to physicians' professional performance in the UK.
Nurses' Assessment	{1,2,3,4,5} 1 = very poor 5 = very good	
HR Assessment	{1,2,3,4,5} 1 = very poor 5 = very good	
Uses effective and appropriate clinical problem-solving skills	{1,2,3,4,5} 1 = very poor 5 = very good	(Ho and Baker, 2012) used Knowledge, skills, performance, safety and quality, communication, partnership, teamwork, and maintaining trust.
Understand the limits of one's Knowledge and expertise	{1,2,3,4,5} 1 = very poor 5 = very good	
The ability to evaluate the treatment methods used.	{1,2,3,4,5} 1 = very poor 5 = very good	
The ability to plan and develop work.	{1,2,3,4,5} 1 = very poor 5 = very good	
Skill in diagnosing the disease and determining the appropriate treatment.	{1,2,3,4,5} 1 = very poor 5 = very good	
Follow up on developments in the field of specialization.	{1,2,3,4,5} 1 = very poor 5 = very good	
Familiarity with work systems and procedures.	{1,2,3,4,5} 1 = very poor 5 = very good	
Follow-up of laboratory results and medical examinations.	{1,2,3,4,5} 1 = very poor 5 = very good	
Carrying out the requirements of emergency shifts.	{1,2,3,4,5} 1 = very poor 5 = very good	
Uses information technology to optimize patient care	{1,2,3,4,5} 1 = very poor 5 = very good	
Acknowledges medical errors	{1,2,3,4,5} 1 = very poor 5 = very good	
Develops systems to reduce medical errors	{1,2,3,4,5} 1 = very poor 5 = very good	
Ensures quality	{1,2,3,4,5} 1 = very poor 5 = very good	

Number of Publications	Range	**(**Kuemmerle**, 2014)** added standards to the Maintenance of Certification (MOC) program, which focuses on professional standing as it contributes to improving medical practice.
Number of Research Projects	Range	
Number of Research Presentations	Range	
Medical error rate	Range	
Number of complaints	Range	I proposed a cognitive model that classifies medical errors based on the individual's level and the extent of his interaction with technology.
The number of successful operations	Range	
The number of unsuccessful operations	Range	

Table 1 demonstrates the different set of features, their type, and why they were chosen. The first column represents the set of features used to evaluate the physician. The second column (Data description) has a description of each feature. For example, the ID number contains unique numerical data. Each record has a different key value because it will be the primary key to accessing the physician's data. As for the gender and nationality features, it will be a multiple choice so that the gender will be (male/female) and the nationality (citizen/ Resident), and it is preferable that the data be in age as a range. For example, ages 25-34 represent the number 1. From 35 to 44, the number 2 is described, and so on. These features were chosen in the personal data based on (Campbell et al, 2008) studies that added gender and age in one of the evaluation questionnaires. This paper indicates that nationality must be added to performance evaluation since the contracts and assessments of foreign physicians may differ from the contracts of national physicians in some countries. The academic qualifications, job title, and department were listed and set as multiple-choice. The years of experience were set as a range, i.e., less than 5 years, we denote the number 1; between 5-10 years, we denote the number 2; and between 10-15 years, we denote the number 3, and so on. Academic qualifications are considered one of the essential features in evaluating performance. The more qualified the physician is, the more likely he is to receive a good evaluation. As add (Overeem et al, 2012) practical experience and job title also play a role in increasing the physician's performance rate. The study added a department feature to know which department is most exposed to medical errors. The study also added an advantage to physicians affiliated with the Health Authority for Specialties or the Ministry of Health due to the difference in the evaluation mechanism between them and the type of data (yes/no). You're understanding and application of these features are crucial to the success of the physician evaluation process, highlighting your value and integral role in this process.

As for other features (Patient Assessment, Nurses Assessment, and HR Assessment, Using effective and appropriate clinical problem-solving skills, Understanding the limits of one's knowledge and expertise, evaluating the treatment methods used, The ability to plan and develop work, Skill in diagnosing the disease and determining the appropriate treatment, Following up on developments in the field of specialization, Familiarity with work systems and procedures, Follow-up of laboratory results and medical examinations, Carrying out the requirements of emergency shifts, Utilizing information technology to optimize patient care, Acknowledging medical errors, developing systems to reduce medical errors, and Ensuring quality). The data type for all these features was chosen to be numeric, where 1 = very poor and 5 = very good. Ho and Baker (2012) relied on questionnaires from patients and colleagues to collect data related to the physician's performance, and Kuemmerle (2014) used knowledge and skills to evaluate performance. Therefore, the study suggests some features of internal and external evaluations to obtain a more accurate assessment. The study also proposes more features that demonstrate the physician's level of medical knowledge.

There was also another group of features, which were: (the number of Publications, Number of Research Projects, Number of Research Presentations, Medical error rate, Number of complaints, the number of successful operations, the number of unsuccessful operations). The type of these features was chosen to be on a range from 1-5. These features were selected based on studies that focused on professional standing to improve medical practice. Therefore, the study proposes evaluating physician research projects and presentations that could contribute to improving physician performance. The study also suggests including the rate of medical errors, the number of complaints, the number of successful operations, and the number of failed operations in evaluating the physician's performance. The goal of this study is to provide a comprehensive method for assessing physician performance in medical practice. Zhang et al. (2004) proposed a cognitive model that classifies medical errors based on the individual's level and the extent of his interaction with technology. Based on this table, this paper proposes to place all the proposed features for evaluating physicians' performance in a structural chart within nine classifications, as shown in Figure 2.

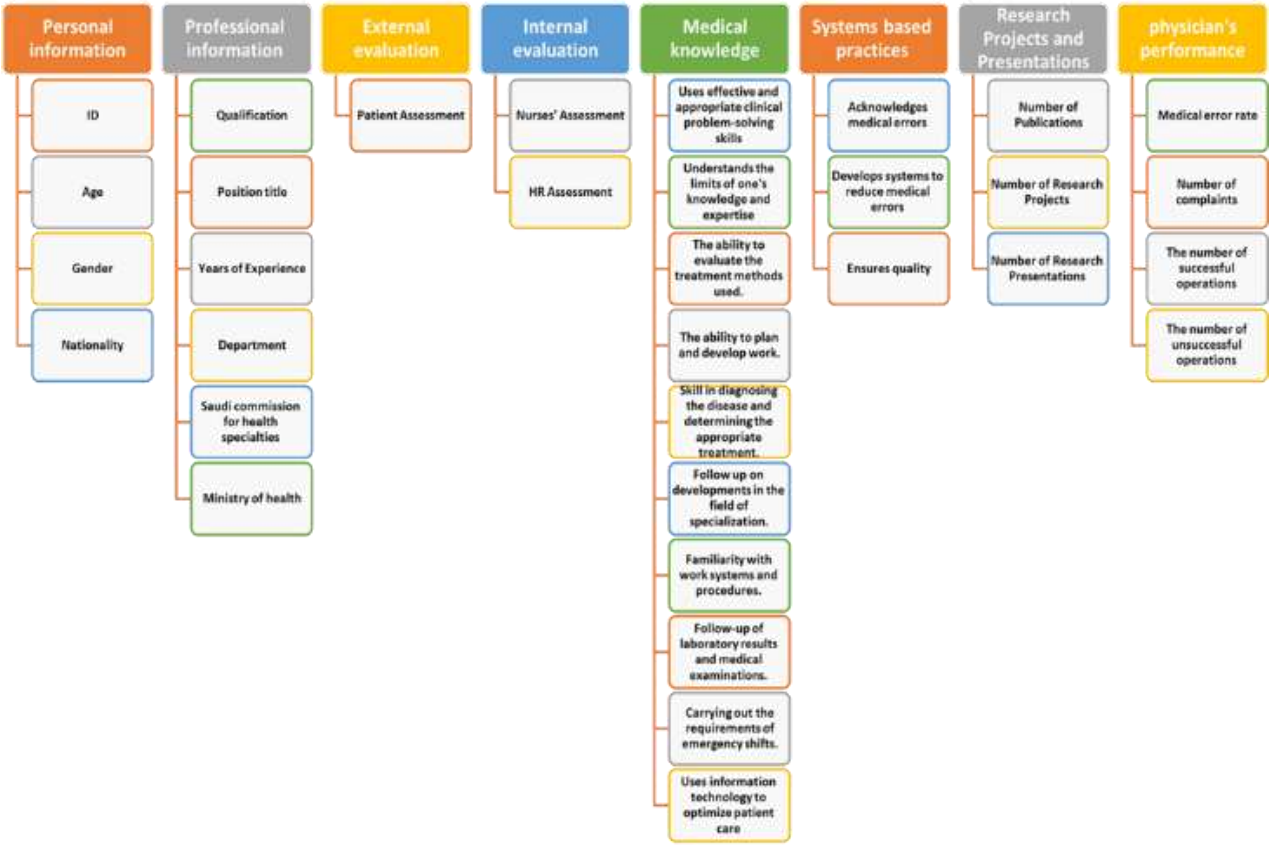

Figure 2. Proposed features in evaluating physicians' performance

Figure 2 showcases a meticulous classification of features into nine groups, including Personal Information, Professional Information, External Evaluation, Internal Evaluation, Medical Knowledge, Systems Practices, Research Projects and Presentations, and Physician's Performance. This thorough classification ensures that all essential aspects of healthcare professionals' performance evaluation are covered. Each classification contains a set of features, such as personal information, which includes several features (ID, Age, Gender, and Nationality). Professional information features include (Qualification, Position title, Years of Experience, Department, Saudi Commission for Health Specialties, and Ministry of Health). External evaluation shows (Patient Assessment), while internal evaluation includes (Nurses' Assessment and HR Assessment). Medical knowledge contains (Using effective and appropriate clinical problem-solving skills, Understanding the limits of one's knowledge and expertise, evaluating the treatment methods used, The ability to plan and develop work, Skill in diagnosing the disease and determining the appropriate treatment, Following up on developments in the field of specialization, Familiarity with work systems and procedures, Follow-up of laboratory results and medical examinations, Carrying out the requirements of emergency shifts, Uses information technology to optimize patient care). Systems-based practices indicate (Acknowledging medical errors, developing systems to reduce medical errors, and Ensuring quality). The classification of Research Projects and Presentations included a group of features, including (the number of Publications, Number of Research Projects, Number of Research Presentations), Physician's performance (Medical error rate, Number of complaints, the number of successful operations, the number of unsuccessful operations).

4. EXPERIMENT SETUP

Data were collected from an Egyptian hospital based on the features identified in this study, as the hospital relies on manual evaluation. The hospital's data was collected based on the features prepared in this study, and then data was collected about the Physicians, including their qualifications and experience. However, it was noted that some data was not available in this hospital or was not applicable. The dataset, which included information on physicians and consisted of 500 records, was used by many entities, including human resources management, department heads, and patients, to evaluate the physicians. The dataset contains 34 features of information related to the physician. Still, data on some features were not available in the hospital due to the differing mechanisms and nature of hospital evaluation across countries. For instance, in Egyptian hospitals, the Saudi Commission for Health Specialties is not involved. Additionally, some features did not have identical information for all physicians, adding to the complexity of the evaluation process. Figure 3 illustrates the steps involved in applying regression analysis and discovering influential features, highlighting the challenges in evaluating physicians.

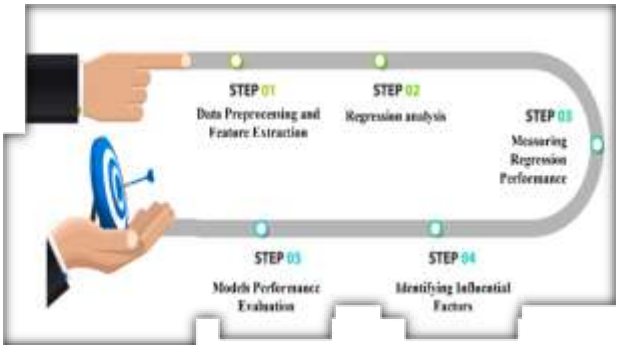

Figure 3. Regression analysis steps

4.1 Data Preprocessing and Feature Extraction

The following filtration procedures were meticulously performed, ensuring the efficiency and accuracy of the analysis:

- Unique identifier (one feature): The ID feature, while unique to each physician, was excluded as it does not contribute to the analysis or modelling. Its exclusion was based on the understanding that the uniqueness of the ID does not provide any additional information for the analysis. Unavailable data (4 features): Attributes related to physicians being part of 'The Ministry of Health' or 'The Saudi Commission for Health Specialties' were excluded because they are not applicable in Egyptian hospitals. Similarly, the 'Number of Research Projects' and 'Number of Research Presentations' features were also excluded because they have empty data.
- Identical values (14 features): some features contain data that is identical to all Physicians, including 'Qualification,' ten 'Medical knowledge' attributes, and three 'Systems-based practices' features since all physicians hold PhD degrees, acquire all required knowledge, and perform the same practices. Moreover, some features required further processing, as follows.
- Operations (2 features): The medical error rate is calculated based on two attributes: 'The number of successful operations' and 'The number of unsuccessful operations. Having these features in the model can distort regression as they are perfectly correlated. Hence, these two features are merged into one feature representing the 'Total number of operations' and then excluded from the data.
- Age: the ages of physicians have been categorized into several groups based on age ranges to help generalize the findings, as follows:
  40-49 years old: Group 1
  50-59 years old: Group 2
  60-69 years old: Group 3
  70-79 years old: Group 4
- Position title: the data contains information about the medical department as part of the position titles of the physicians. For example, one Physician has the position title of 'Vascular surgery consultant,' and another physician has the position title of 'Neurosurgery Consultant.' However, these two titles refer to the same role, which is 'Consultant.' Therefore, the 'Position title' feature is processed to contain only the role(s) of each physician by eliminating any information related to department or university names. Figure 4 shows the proposed features after processing.

ID	Age	Age Groups	Gender	Nationality	Department	Position title	Years of Experie	Patient Assessm	Nurses' Assessm	HR Assessment	Supervisor Asse	Number of Publi	Medical error rat	Number of comp	Total number of
1.0	59	2	M	Egyptian	General surgery	Head/Professor	34	5.0	5.0	4.9	4.9	209	0.02	0.01	1204
2.0	52	2	F	Egyptian	General surgery	Assistant Head/I	27	4.3	4.6	5.0	4.7	136	0.02	0.04	904
3.0	58	2	M	Egyptian	General surgery	Professor/Consu	33	4.0	4.7	4.9	4.9	135	0.06	0.01	812
4.0	67	3	M	Egyptian	General surgery	Professor/Consu	42	3.9	4.8	4.8	5.0	99	0.05	0.03	651
5.0	66	3	M	Egyptian	General surgery	Professor/Consu	41	4.2	4.5	5.0	4.8	168	0.04	0.03	700
6.0	66	3	M	Egyptian	General surgery	Consultant	41	4.8	4.6	4.7	4.9	22	0.03	0.01	924
7.0	64	3	M	Egyptian	General surgery	Professor/Consu	36	4.0	4.7	4.7	4.9	188	0.02	0.03	1141
8.0	64	3	M	Ukrainian	General surgery	Professor/Consu	35	4.8	4.7	4.7	4.9	78	0.06	0.02	358
9.0	63	3	M	Egyptian	General surgery	Consultant	38	4.5	5.0	4.8	4.9	31	0.04	0.04	584
10.0	63	3	F	Egyptian	General surgery	Professor/Consu	38	4.5	4.8	4.9	4.9	114	0.02	0.02	940
11.0	62	3	F	Egyptian	General surgery	Professor/Consu	37	4.9	5.0	4.9	4.8	76	0.05	0.03	688
12.0	61	3	M	Egyptian	General surgery	Professor/Consu	36	4.3	4.6	4.8	5.0	74	0.05	0.02	765
13.0	59	2	M	Egyptian	General surgery	Professor/Consu	34	4.3	4.7	4.7	4.8	106	0.02	0.04	1068
14.0	58	2	M	Egyptian	General surgery	Professor/Consu	33	5.0	4.6	4.7	4.7	136	0.04	0.03	806
15.0	58	2	M	Egyptian	General surgery	Consultant	33	4.2	4.5	4.7	4.8	14	0.03	0.03	655
16.0	58	2	M	Egyptian	General surgery	Consultant	33	4.9	4.8	5.0	4.9	31	0.03	0.03	730
17.0	57	2	M	Egyptian	General surgery	Consultant	32	4.9	4.9	4.9	4.7	45	0.04	0.03	384
18.0	56	2	M	Egyptian	General surgery	Consultant	31	5.0	4.5	4.7	4.8	19	0.04	0.04	1026
19.0	56	2	M	Egyptian	General surgery	Consultant	31	4.7	4.9	4.7	4.7	33	0.03	0.02	980
20.0	55	2	F	Egyptian	General surgery	Consultant	30	4.0	4.5	4.8	4.9	62	0.02	0.02	1126
21.0	53	2	F	RUSSIAN	General surgery	Consultant	28	4.6	4.7	4.9	4.8	32	0.06	0.02	818
22.0	52	2	M	Egyptian	General surgery	Consultant	27	4.5	4.5	4.6	4.9	41	0.04	0.01	426
23.0	49	1	F	Egyptian	General surgery	Professor/Consu	24	4.8	4.4	4.8	4.9	81	0.04	0.02	631
24.0	48	1	M	Egyptian	General surgery	Professor/Consu	23	4.5	4.6	4.0	4.0	166	0.03	0.03	1218
25.0	49	1	M	Egyptian	General surgery	Professor/Consu	24	4.3	4.8	4.9	4.9	115	0.02	0.02	420
26.0	61	3	M	Egyptian	General surgery	Professor/Consu	36	3.9	4.6	4.9	4.7	48	0.06	0.02	713

Figure 4. The proposed data set for evaluating physicians' performance

After processing the data, extracting essential features, and deleting uproarious and identical data, 16 features were found that can be used to evaluate physicians' performance. In the next step, the regression analysis process will be applied using data mining technology to discover the most important features affecting the physician's evaluation.

4.2 Regression Analysis

The study proposes applying a data mining technique using regression analysis Where Draper & Smith (1998) mentioned in his study that applying regression analysis helps to determine the most influential factors. The study utilizes the power of Python scripts to conduct this analysis, showcasing the use of modern tools in healthcare research. Regression analysis is a powerful statistical technique used to model and analyse relationships between a dependent variable (objective) and one or more independent variables (traits or features). The importance of regression analysis is that it allows for measuring various factors, such as professional standards and results. Therefore, it will enable healthcare institutions to implement targeted interventions to reduce the rate of medical errors and improve patient safety, potentially transforming healthcare outcomes. The study suggests the use of three distinct regression models, as follows.

- Linear Regression (Groß, 2003): is a practical tool for determining which factors are statistically significant predictors of the medical error rate. It provides transparent and interpretable coefficients and significance levels for each feature, making it a reliable method for healthcare research.
- Lasso Regression (Tibshirani, 1996): is a versatile tool for dealing with overlapping and multicollinearity in data. By reducing the coefficients of less important features to zero, it keeps only the most important features in the model, making it adaptable to different healthcare settings.
- Ridge Regression (Hoerl & Kennard, 1970): provides an understanding of feature importance, especially in datasets where many features are correlated, helping to understand their collective impact on medical error rates.

4.3 Measuring Regression Performance

The performance of regression models is measured using several key metrics. In this study, we use the most common metrics: R-squared (R^2) and Root Mean Squared Error (RMSE).

- R-squared (R^2) In (Lewis-Beck and Skalaban, 1990) study, plays a significant role in our study, as it represents the proportion of variance in the dependent variable that is predictable from the independent variables.
- A higher R^2 indicates a better fit. Root Mean Squared Error (RMSE) (Chai & Draxler, 2014) add, is a crucial metric in our study, as it measures the average of the squares of the errors between actual and predicted values. Lower values indicate a better fit. For this study, we have implemented the widely accepted 80-20 split ratio.

This approach, which dedicates 80% of the dataset to model training and reserves 20% for testing, is a robust method that installs confidence in the evaluation of model performance. It ensures the model has sufficient data to learn effectively while retaining a substantial portion for rigorous testing.

4.4 Identifying Influential Factors

Identifying influential factors is crucial for each regression model used. These factors play a significant role in the model's predictive power and our understanding of the underlying relationships.

- P-values (Harvey & Sotardi, 2018): (linear regression): These values precisely measure the statistical significance of each feature's coefficient. A lower p-value indicates that a feature is likely to have a meaningful relationship with the dependent variable, independent of other features.
- Coefficient Values definition of (Fraser et al,1999) (Linear Regression, Lasso, Ridge): The size and sign of the coefficients indicate the direction and magnitude of each feature's influence on the dependent variable. In Lasso regression, coefficients of some variables may be shrunk to zero, directly indicating their non-importance.

Based on the above values, the significance of features was annotated with stars ('', '', ''). Signs ('+', '–') were added to indicate the direction of the relationship for Lasso and Ridge's models. These signs make it clear whether the relationship is direct or inverse, providing a clear understanding of the influence.

4.5 Models Performance Evaluation

Using the three regression models, the model should examine the R^2 Score and RMSE for predicting the 'Medical error rate'. Table 2 presents a Summary of the performance of the three models based on the R^2 Score and RMSE.

Table 2. Summary of the performance of the three models based on the R^2 Score and RMSE

Model	R^2 Score	RMSE
Linear Regression	0.242	0.01436
Lasso	0.413	0.01264
Ridge	0.397	0.01281

Table 2 highlights a significant finding: the Linear Regression model, with its R^2 Score of 0.242 and the highest RMSE value of 0.01436, is limited in its ability to explain the variance in the medical error rate. This suggests that its predictions are less accurate on average than the other models. The model's simplicity and its inability to capture nonlinear relationships and interactions among features may be the contributing factors to its performance.

Conversely, the Lasso and Ridge regression models exhibit significant performance improvements. Lasso achieves an R^2 Score of 0.413 and an RMSE of 0.01264, while Ridge achieves a slightly lower R^2 Score of 0.397 and an RMSE of 0.01281. The enhanced performance of these models can be attributed to their incorporation of regularization terms, which help mitigate overfitting and allow the models to generalize better to unseen data. The Ridge model shows promise with its slightly lower R^2 Score but comparable RMSE, indicating its potential to perform well in predicting medical error rates. Table 3 shows the results of feature importance in predicting the medical error rate using the three regression models. Summarized as follows.

Table 3. Importance of features in predicting the medical error rate using the three models.

Attribute/Model	Lasso Regression	Ridge Regression	Linear Regression
Nationality	(+) ***	(+) ***	(+) *
Position title	(-) ***	(-) ***	(-) **
Years of Experience	(+) **	(+) **	
Number of Publications	(+) *	(+) *	
Number of operations	(-) *	(-) *	(-) ***
Age Groups		(-) ***	
Gender		(-) *	
Department		(+) ***	
Patient Assessment		(+) *	
Nurses' Assessment		(-) *	
HR Assessment		(-) ***	
Supervisor Assessment		(-) *	
Number of complaints		(-) *	

Nationality and Position Title stand out with high significance across Lasso and Ridge regression models, indicating a strong and consistent impact on the medical error rate. However, the relevance of the Nationality attribute could be influenced by the fact that about 95% of the Physicians in our dataset are Egyptians. Also, it was revealed that 'Ukrainians' and 'Syrians' are more likely to have higher error rates than others, whereas 'Germans' are more likely to have lower error rates than others. Cultural and language differences may play a role in medical errors. Concerning position titles, the study found that Physicians whose titles are within the combinations of 'Assistant Head,' 'Professor,' or 'Consultant' are associated with higher error rates than those of other ranks. This could indicate that Physicians who are involved in other administrative or teaching/research responsibilities may be better than those who are more practitioners in the field, highlighting the need for further investigation into the impact of administrative and teaching duties on medical error rates. The findings from years of experience and publications recognized by both Lasso and Ridge regressions present an intriguing twist.

Contrary to expectations, these attributes are positively associated with higher error rates, challenging the common belief that more experience and research publications indicate successful operations. This unexpected result invites further exploration and discussion. The finding on the number of operations, with a significant inverse association in Lasso, Ridge, and Linear Regression models, carries practical implications. Unlike years of experience, these results suggest that the more operations a Physician performs, the lower the rate of medical errors. This insight equips hospitals with a clear understanding that the number of operations a Physician performs, regardless of their long employment history, is a crucial factor in their success. Features such as Age Groups, Gender, Department, Patient Assessment, Nurses' Assessment, HR Assessment, and Supervisor Assessment show varied levels of importance and significance across the models, indicating their different contributions to the medical error rate. Notably, the Department attribute shows of high significance in Ridge regression models, but it needs to be more emphasized in other models. Physicians from the 'Neurosurgery' department have higher error rates than any other department, which is justified by how complex surgeries of this kind can be. In addition, the HR Assessment showed higher importance than other assessments.

Through the study, the Lasso model revealed that the factors that can influence the evaluation of Physicians are, respectively, in order of importance: nationality, job title, years of experience, number of publication papers, and finally, the number of operations performed by the Physician, while the linear regression model showed that the influential factors are, respectively: number of operations and title, “Career and nationality.” On the other hand, the Ridge model indicated 13 characteristics that would influence the evaluation of Physicians, which are in order of importance: nationality, job title, age, department, and human resources evaluation. After that, years of experience come into the second stage of importance, and finally, the number of publication papers, number of operations, gender, patient evaluation and evaluation, nurses, supervisor, and number of complaints. Therefore, the study proposes to take the crucial features from the three models and apply data mining operations to them, including clustering and classification, with the potential to revolutionize our understanding of physician evaluation and improve healthcare outcomes.

5. CONCLUSION

In this Paper, many challenges were faced, which were represented by the lack of studies that use data mining in evaluating the performance of the physician, as well as the lack of a ready data set for physicians in hospitals, which necessitated the creation of a data set from scratch based on the proposed features. Most of the proposed features did not achieve the desired benefit, as the physicians in the hospital have the same qualifications, and there were many identical values and empty values. In practical terms, the regression analysis provides healthcare administrators with a data-driven roadmap to reduce medical error rates. The significance of features in predicting medical errors suggests areas where targeted interventions, such as culturally sensitive training or reevaluating the roles and responsibilities of higher-ranking medical staff, could be beneficial.

As key decision-makers, healthcare administrators play a crucial role in implementing these interventions, thereby addressing underlying systemic issues such as differing educational backgrounds or the impact of administrative duties on patient care. Besides, the unexpected finding related to specific attributes, such as years of experience and number of publications being correlated with higher error rates, implies that ongoing professional development and active clinical practice are crucial for maintaining high standards of care.

This insight can guide hospitals in designing continuous education programs and encourage a culture of regular hands-on practice, regardless of a physician's tenure or academic achievements. This is also supported by the inverse relationship between the number of operations and error rates, which reinforces the importance of practical experience over theoretical knowledge, thus encouraging hospitals to prioritize allocating surgeries to enhance physicians' skills and reduce errors. Overall, these insights empower healthcare administrators to refine recruitment, training, and evaluation processes, thereby improving patient safety and care quality.

6. FUTURE WORK

This paper referred to the personal information and individual skills of physicians in assessing physicians' performance, in addition to evaluating them from several different points of view (patient evaluation, supervisor evaluation, and human resources evaluation) as the study applied a framework that would contribute to discovering the features on which physicians in hospitals are evaluated. Therefore, the paper proposes that in future work, a combination of clustering and classification techniques be applied to group physicians into groups with similar characteristics using the K-means algorithm. Analysing each group will indicate the appropriate decision to help human resources improve the physician's performance. Next, a classification technique is applied to help HR predict the appropriate decision previously identified in the clustering step for new physicians. This prediction will be based on the specific characteristics of the study, such as the physician's experience, patient feedback, and adherence to clinical guidelines. The study indicated some areas of improvement that could be applied in the future. It is possible to apply this framework to a hospital database using the features described in the survey to obtain accurate results.

References:

Bindels, E., Boerebach, B., Scheepers, R., Nooteboom, A., Scherpbier, A., Heeneman, S., & Lombarts, K. (2021). Designing a system for performance appraisal: Balancing physicians' accountability and professional development. *BMC Health Services Research, 21*(1), 1–12. https://doi.org/10.1186/s12913-020-06017-5

Brennan, N., Bryce, M., Pearson, M., Wong, G., Cooper, C., & Archer, J. (2014). Understanding how appraisal of doctors produces its effects: A realist review protocol. *BMJ Open, 4*(6), e005466. https://doi.org/10.1136/bmjopen-2014-005466

Campbell, J., Richards, S., Dickens, A., Greco, M., Narayanan, A., & Brearley, S. (2008). Assessing the professional performance of UK doctors: An evaluation of the utility of the general medical council patient and colleague questionnaires. *BMJ Quality & Safety, 17*(3), 187–193. https://doi.org/10.1136/qshc.2006.021816

Cassel, C. K., & Jain, S. H. (2012). Assessing individual physician performance: Does measurement suppress motivation? *JAMA, 307*(24), 2595–2596. https://doi.org/10.1001/jama.2012.5820

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? *Geoscientific Model Development Discussions, 7*(1), 1525–1534. https://doi.org/10.5194/gmdd-7-1525-2014

Cola, P. A., & Wang, Y. (2017). Discovering factors that influence physician scientist success in academic medical centers. *Academy of Management Proceedings, 2017*(1), 12714. https://doi.org/10.5465/ambpp.2017.12714abstract

Dias, R. D., Gupta, A., & Yule, S. J. (2019). Using machine learning to assess physician competence: A systematic review. *Academic Medicine, 94*(3), 427–439. https://doi.org/10.1097/ACM.0000000000002540

Draper, N. R., & Smith, H. (1998). *Applied regression analysis* (Vol. 326). John Wiley & Sons. https://www.wiley.com/en-us/Applied+Regression+Analysis%2C+3rd+Edition-p-9780471170822

Eva, K. W. (2018). Cognitive influences on complex performance assessment: Lessons from the interplay between medicine and psychology. *Journal of Applied Research in Memory and Cognition*, *7*(2), 177–188. https://doi.org/10.1016/j.jarmac.2018.04.004

Fraser, D. A. S., Wong, A., & Wu, J. (1999). Regression analysis, nonlinear or nonnormal: Simple and accurate p-values from likelihood analysis. *Journal of the American Statistical Association, 94*(448), 1286–1294. https://doi.org/10.1080/01621459.1999.10473877

Groß, J. (2003). *Linear regression* (Lecture Notes in Statistics, Vol. 175). Springer Science & Business Media. https://doi.org/10.1007/978-3-642-55864-1

Harvey, H. B., & Sotardi, S. T. (2018). The Pareto principle. *Journal of the American College of Radiology, 15*(6), 931. https://doi.org/10.1016/j.jacr.2017.12.032

Ho, T. K., & Baker, D. M. (2012). Appraisal and revalidation. *Surgery (Oxford), 30*(9), 447–454. https://doi.org/10.1016/j.mpsur.2012.07.004

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics, 12*(1), 55–67. https://doi.org/10.1080/00401706.1970.10488634

Jin, H., Zhou, J., Zhang, J., & Fu, Y. (2024). Factors influencing healthcare workers' performance before and after the coronavirus disease 2019 pandemic: A bibliometric analysis with supplementary comparative analysis. *Work*, *Preprint*, 1–20. https://doi.org/10.3233/WOR-213580

Kuemmerle, J. F. (2014). ABIM maintenance of certification 2014: Navigating the challenges to find opportunities for success. *Gastroenterology, 147*(2), 260–263. https://doi.org/10.1053/j.gastro.2014.06.008

Lewis-Beck, M. S., & Skalaban, A. (1990). The R-squared: Some straight talk. *Political Analysis*, *2*, 153–171. https://doi.org/10.1093/pan/2.1.153

Mirfat, S., Azzuhri, M., & Hakim, L. (2018). Physician performance analysis based on individual, psychological and organizational factors in medical resume filling. *Journal of Applied Management (JAM), 16*(3), 520–529. https://doi.org/10.21776/ub.jam.2018.016.03.17

Edo, G. I., Nwosu, L. C., Yesiltas, M., Agoh, E., & Lawal, R. A. (2023). Evaluation of factors influencing physician–patient communication in healthcare service delivery. *BOHR International Journal of General and Internal Medicine, 2*(1), 27–33. https://doi.org/10.54646/bijgim.2023.15

Overeem, K., Wollersheim, H. C., Arah, O. A., Cruijsberg, J. K., Grol, R. P., & Lombarts, K. M. (2012). Evaluation of physicians' professional performance: An iterative development and validation study of multisource feedback instruments. *BMC Health Services Research, 12*(1), 1–11. https://doi.org/10.1186/1472-6963-12-80

Sargeant, J. (2019). Future research in feedback: How to use feedback and coaching conversations in a way that supports development of the individual as a self-directed learner and resilient professional. *Academic Medicine, 94*(11S), S9–S10. https://doi.org/10.1097/ACM.0000000000002871

Shi, Y., Li, P., Yu, X., Wang, H., & Niu, L. (2018). Evaluating doctor performance: Ordinal regression-based approach. *Journal of Medical Internet Research, 20*(7), e240. https://doi.org/10.2196/jmir.8826

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society*: Series B (Methodological*), 58*(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

William, W., Kholil, K., Sukwika, T., & Ariyani, N. (2022). Analysis of factors affecting the performance of medical lecturers: Case study at private "X" University Indonesia. *Dinasti International Journal of Management Science, 3*(6), 1130–1145. https://doi.org/10.31933/dijms.v3i6.1248

Zhang, J., Patel, V. L., Johnson, T. R., & Shortliffe, E. H. (2004). A cognitive taxonomy of medical errors. *Journal of Biomedical Informatics, 37*(3), 193–204. https://doi.org/10.1016/j.jbi.2004.04.004

Amani Ghazzawi
Department of Information Systems, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah.
Department of Management Information Systems, Faculty of Business Administration, Taif University, Taif
Saudi Arabia
amani.ghazzawi@yahoo.com
ORCID 0009-0007-6591-7513

Alaa Almagrabi
Department of Information Systems, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah
Saudi Arabia
aalmagrabi3@kau.edu.sa
ORCID 0000-0002-4858-9366

Hanaa Namankani
Department of Information Systems, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah
Saudi Arabia
hnamankani@kau.edu.sa
ORCID 0000-0003-4536-3238