Engineering Applications of Artificial Intelligence 133 (2024) 108610

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

On the capacity of artificial intelligence techniques and statistical methods to deal with low-quality data in medical supply chain environments

Francisco Javier Santos Arteaga [a,*], Debora Di Caprio [b], Madjid Tavana [c,d], David Cucchiari [e,f], Josep M. Campistol [e,f,g,h], Federico Oppenheimer [e,f,g], Fritz Diekmann [e,f,g,h], Ignacio Revuelta [e,f,g,h,**]

[a] *Departamento de Análisis Económico y Economía Cuantitativa Universidad Complutense de Madrid, Madrid, Spain*
[b] *Department of Economics and Management University of Trento, Trento, Italy*
[c] *Business Systems and Analytics Department, Distinguished Chair of Business Analytics, La Salle University, Philadelphia, USA*
[d] *Business Information Systems Department, Faculty of Business Administration and Economics, University of Paderborn, Paderborn, Germany*
[e] *Renal Transplant Unit. Department of Nephrology and Kidney Transplant. Hospital Clinic, Barcelona, Spain*
[f] *Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain*
[g] *National Network for Kidney Research (REDinREN), Carlos III Royal Institute. Ministry of Health, Spain*
[h] *University of Barcelona*

ARTICLE INFO

Keywords:
Information retrieval
Data quality
Artificial neural networks
Regression analysis
Supply chains
Kidney transplantation

ABSTRACT

We illustrate the capacity of Artificial Intelligence (AI) and Machine Learning (ML) techniques to preserve consistent categorization abilities whenever the quality of the data decreases, displaying mistakes or mismatches across matrix entries, while standard statistical methods exhibit significant modifications in the value of the corresponding coefficients. We design algorithms of different complexity to generate a series of comparable profiles. These profiles are compared within environments that allow for an immediate identification of the generating algorithms and within increasingly complex settings involving almost identical profiles derived from different algorithms. AI and ML techniques outperform standard statistical methods when distinguishing the algorithms generating the profiles. Building on these results, we perform a retrospective analysis where AI and ML techniques are applied to two empirical scenarios defined by different data series of patients transplanted through the period 2006–2019. The first scenario contains the variables describing the evolution of patients inputted correctly. In the second, we modify the content of the vectors of characteristics defining the evolution of patients by exchanging the values of a subset of realizations from two categorical variables. AI and ML techniques are consistently accurate when categorizing patients correctly within both scenarios, a feature particularly relevant when the quality of the information sources composing the medical chain varies. This latter problem is exacerbated among hospitals located in developing countries, where the quality of the data gathered limits their identification and extrapolation capacities.

1. Introduction

Physicians display a tendency to distrust the results obtained from the implementation of artificial intelligence (AI) and machine learning (ML) techniques (Bae, S. et al., 2020; Wynants et al., 2020; Lancet. Artificial intelligence for COVID-19, 2021). Despite this fact, these techniques have been consistently applied to identify potential patterns within the current COVID-19 crisis (Arora et al., 2020; Massie et al., 2020; Rasheed et al., 2020; Vaishya et al., 2020). Classification problems consist of a set of predictors, that is, features describing the different alternatives, defined via n-dimensional vectors and an outcome per alternative, namely, the class to which the alternative belongs. AI and ML techniques evaluate the features of each alternative together with its class and learn from them so that whenever an alternative is observed, the class to which it belongs can be predicted. The black-box quality of the processes performed by these techniques, as opposed to the more

* Corresponding author.
** Corresponding author. Department of Nephrology and Renal Transplantation, Hospital Clínic, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Villarroel 170 (Escala 10 – Planta 5), E-08036 Barcelona, Spain.
E-mail addresses: fransant@ucm.es (F.J. Santos Arteaga), irevuelt@clinic.cat (I. Revuelta).

https://doi.org/10.1016/j.engappai.2024.108610
Received 29 October 2023; Received in revised form 23 April 2024; Accepted 9 May 2024
Available online 28 May 2024

intuitive statistical methods commonly applied, add to their interpretation difficulties when biasing physicians against their implementation.

The main objective of the current paper is to illustrate how the identification capacities of these techniques are determined and can be enhanced by the way the features are ordered within the vectors defining the alternatives. That is, AI and ML techniques can categorize alternatives even if their characteristic features are presented in a way that prevents their immediate identification using standard statistical techniques. The main purpose of the analysis is highlighting the capacity of AI and ML techniques to preserve a consistent identification capacity whenever the quality of the data decreases, displaying mistakes or mismatches in several matrix entries, while standard statistical techniques would exhibit significant modifications in the value of the corresponding coefficients.

The data used to illustrate the categorization capacities of AI and ML techniques is obtained by simulating the sequential information retrieval processes defined by decision makers (DMs) in online search environments. We are all internet users and perform multiple searches on a daily basis. Thus, the logic on which the different algorithms generating these data are built, and the subsequent retrieval structures, should be sufficiently intuitive. A similar comment applies to the results obtained from the implementation of AI and ML techniques based on the training data produced by the different algorithms.

The online search behavior of users generates raw data describing the queries made and the set of pages clicked. Research on this topic generally aims at extrapolating the utility functions that are consistent with the retrieval behavior observed (Basu, 2018; Victorelli et al., 2020; Sharma et al., 2021). In particular, a variety of research areas have focused on the preferences of DMs determining their information retrieval behavior (Schmitt et al., 2018; Zanganeh and Hariri, 2018; Utku Özmen and Yucel, 2019; Miranda and Miah, 2023).

Nikzad–Khasmakhi et al. (2019) and Sharma et al. (2023) review the literature on expert recommendation systems. Bandit algorithms arise as one of the main techniques used to elicit the preferences of DMs from their linguistic reports and sequential choices (Gabrielli et al., 2024). When applying these models to real-life recommender applications, scholars focus on the constraints both pecuniary and cognitive faced by the actors eliciting the preferences of DMs (Zhao and Yang, 2024). Similarly, when eliciting information from large amounts of data, sampling and noise reduction techniques have been suggested (Jain and Jindal, 2023).

Behavioral data are currently being used to extrapolate the main qualities defining the users so as to enhance the capacity of deep-learning models to predict the subsequent click through rates (Li et al., 2020; Qin et al., 2020). Recent developments concentrate also on the opposite relation, namely, the influence of AI on the behavior of DMs when retrieving information (Verma et al., 2021). All these models leave the complexity of the retrieval process conditioning the behavior of DMs aside, which is indeed the starting point of our formal analysis.

That is, information retrieval process could also be formalized as a decision model where the behavior of DMs is determined by the similarity between their preferred characteristics and those of the alternatives described in the snippets (Di Caprio et al., 2022a, 2022b). Rational DMs have been traditionally considered as a benchmark when formalizing information retrieval behavior (Di Caprio and Santos-Arteaga, 2021a). However, different branches of the literature have highlighted the cognitive limits faced by DMs when evaluating alternatives and retrieving information (Joseph and Gaba, 2020).

The algorithms generated to test the identification capacities of AI and ML techniques define sequential decision processes based on a predetermined retrieval strategy that differs across DMs according to their subjective preferences and information assimilation capacities. These algorithms are based on precise retrieval patterns that become increasingly harder to identify as we approach the maximum number of alternatives that DMs are willing to evaluate. That is, the stochastic structure following from the information retrieval behavior of DMs, which determines the set of retrieval vectors obtained, goes well beyond the simple generation of random numbers used to fill the entries of the vectors.

The current paper assesses the ability of AI and ML techniques to identify the complexity of the information retrieval behavior of DMs. We also illustrate how this capacity can be enhanced by modifying the order of the features composing the retrieval vectors that describe the search processes. In this regard, we must note that the research analyzing the online information retrieval behavior of DMs does not generally account for the distribution of information across the entries composing the data vectors (Li et al., 2020).

1.1. Contribution

The current paper extends the benchmark model introduced by Di Caprio and Santos-Arteaga (2022) by modifying the incentives of DMs to retrieve information and increasing the set of comparisons performed among the retrieval vectors generated. In particular, we design algorithms of different complexity to generate a series of comparable retrieval profiles. Comparisons are performed within straightforward environments – allowing for easily distinguishable profiles –, and increasingly complex settings involving almost identical retrieval vectors derived from entirely different algorithms. In both papers, DMs are categorized according to their capacity to evaluate alternatives, which determines the set of retrieval paths defining the behavior observed. We concentrate on the ten results defining the first page of outcomes delivered by a search engine, which account for most of the retrieval activity of DMs (Dean, 2019).

The main technical contributions of the analysis performed are summarized below.

1. The algorithms are designed to generate online search profiles conditioned by the subjective preferences of DMs as well as their capacity to retrieve information.
2. The algorithms account for all the potential evaluation decisions made by DMs and the subsequent information retrieval paths generated.
3. The distribution of features within the retrieval vectors describing the behavior of DMs conditions the categorization accuracy of ML techniques.
4. ML techniques outperform standard statistical methods when categorizing profiles while facing misprints or errors in the inputs or independent variables.

We apply a battery of AI and ML techniques to demonstrate their ability to distinguish the algorithms generating the profiles across a variety of retrieval environments. Building on these features, we perform a retrospective analysis where AI and ML techniques are applied to two empirical scenarios defined by modifying a data series of 643 patients transplanted through the period 2006–2019. The results derived from the analysis of these scenarios constitute the main contribution of the manuscript from a medical implementation viewpoint.

The first scenario contains the variables describing the evolution of patients inputted correctly. In the second scenario, the content of the vectors of characteristics defining the evolution of patients is modified by exchanging the values of a subset of realizations from two categorical variables. Modifications are applied to a total of 39 patients, 35 of which belong to one output category and 4 to the complementary one. Our main objective is to evaluate the capacity of AI and ML techniques to classify the patients correctly despite the modifications introduced.

Artificial Neural Networks (ANN) as well as the different ML techniques implemented preserve their identification capacities within both scenarios. ANN categorize correctly 81.6% and 82.4% of the patients composing the first and second scenario, respectively. ML techniques display an average accuracy of 80%, with coarse trees (82.6%) and

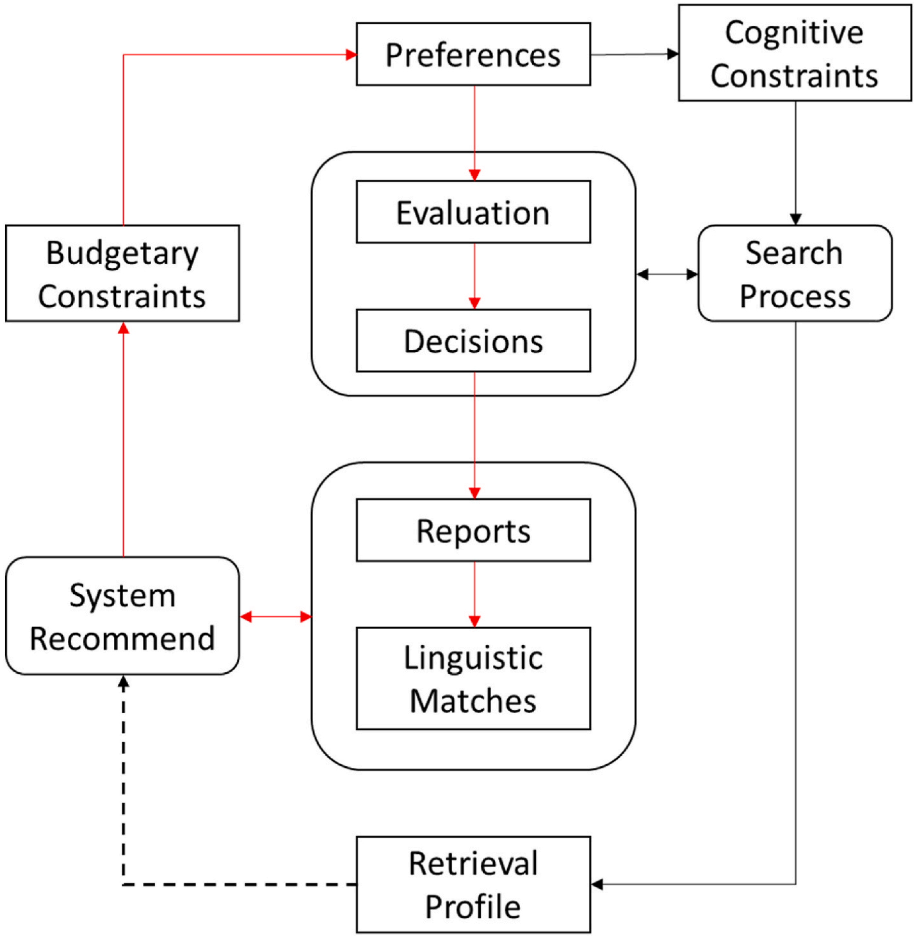


Fig. 1. Information retrieval processes: recommender systems versus sequential search approaches.

linear discriminants (82.3%) exhibiting the largest percentages within the first and second scenario, respectively.

The intuition behind these results follows from the fact that the capacity of AI and ML techniques to identify patterns is determined by the behavior of the whole set of independent variables altogether. The whole set of features, even their relative positioning (Di Caprio and Santos Arteaga, 2022), allows the corresponding techniques to identify the inherent patterns and classify patients correctly. On the other hand, we also illustrate how the independence assumed on the explanatory variables composing a multivariable linear regression model implies that the coefficients are severely affected when shifting values across matrix columns.

From a decision engineering perspective, the algorithms designed to generate the most complex – namely, complete – information retrieval profiles allow to incorporate modifications in the preferences of DMs and simulate the subsequent retrieval scenarios, providing a relevant benchmark for the research performed within the literature on information systems (Zhang et al., 2020; Hong et al., 2021). This quality is highlighted in Fig. 1, which summarizes the main differences between the approach followed by the literature on recommender systems and the current paper. The potential interactions between the algorithms defining the retrieval profiles of DMs and actual recommender systems constitute a potential research line that will be described in the conclusion section.

We conclude by highlighting the complementary qualities of AI and ML techniques relative to standard statistical models. Implementing a statistical model requires defining a structured framework determined by the potential effects derived from a set of independent variables. We emphasize the fact that the results derived from these models can be complemented via AI and ML techniques, particularly in settings facing data quality problems.

Fig. 2 presents a flowchart describing this complementarity while summarizing the evaluation and categorization processes analyzed. Information retrieval processes of different complexity are simulated, and the subsequent profiles paired and categorized using both ML techniques and standard statistical methods. A mirror sequence takes place on the medical side. Patient data have been retrieved to generate profiles whose quality has been purposely modified. Patients are then categorized according to their transplant performance by ML techniques and standard statistical methods and the results compared.

The algorithms have been coded in MATLAB and their outputs used as training inputs within its neural net pattern recognition and clustering apps. The statistical analyses have been performed in SPSS. The code of the algorithms designed to generate the different information retrieval profiles is publicly available from Di Caprio and Santos-Arteaga (2021b).

The paper proceeds as follows. Section 2 reviews the literature dealing with the application of AI and ML techniques to medical environments and healthcare supply chains. Section 3 introduces the retrieval algorithms and the sequential online evaluation frameworks on which they are based. Section 4 defines the medical environment analyzed. Section 5 compares the categorization results obtained from a battery of ML techniques and different standard statistical methods. Section 6 discusses the main consequences derived from the results obtained within both industrial and medical settings. Section 7 concludes and suggests potential extensions.

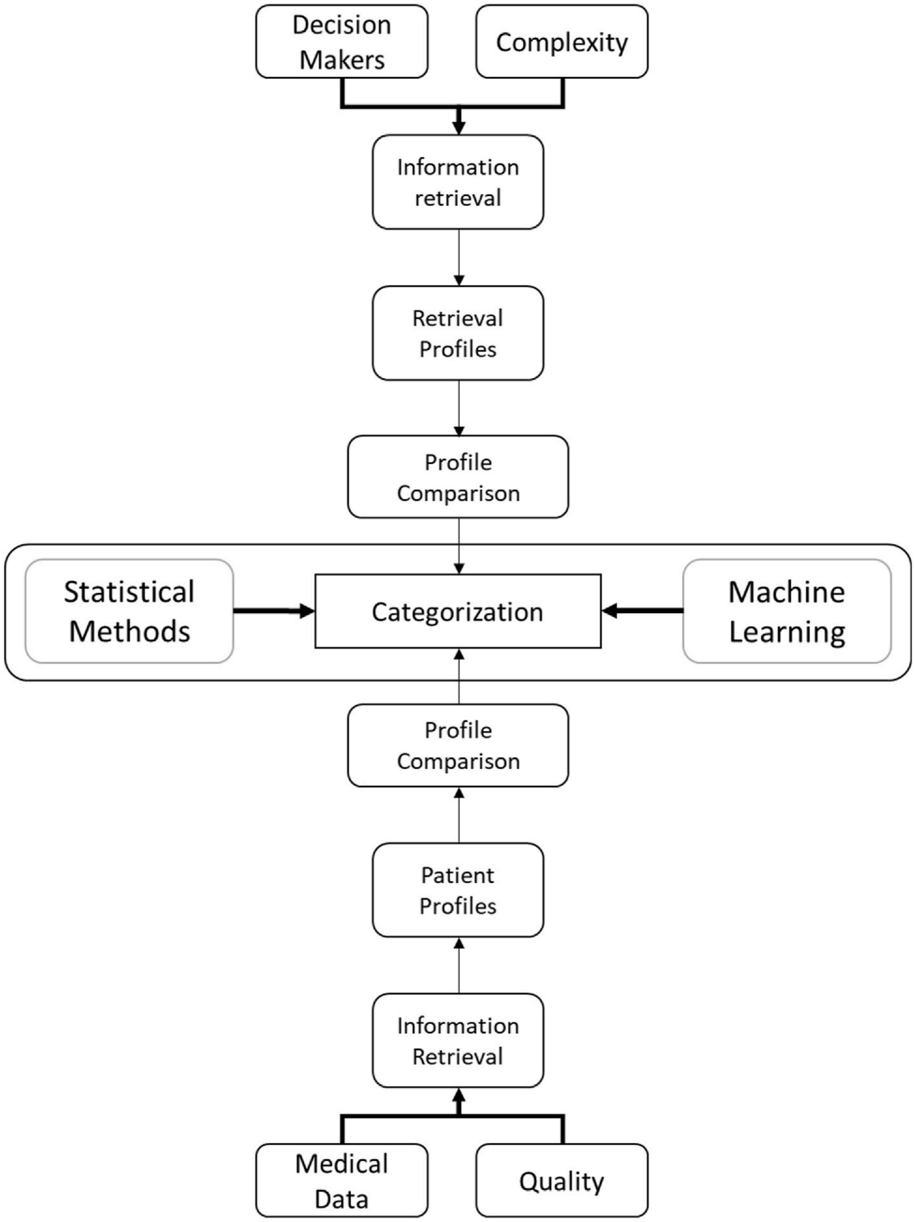


Fig. 2. Evaluation and categorization processes: from information retrieval to medical profiles.

2. Literature review

As stated in the introduction, the application of AI and ML techniques to concrete medical problems remains a problematic issue (Kelly et al., 2019; Matheny et al., 2019; Bajwa et al., 2021). This reticence persists despite the increasing applicability of AI techniques to medical environments (Waring et al., 2020; Jiang et al., 2021; Nwanosike et al., 2022). Patients also tend to distrust the application of AI techniques into the medical domain (Richardson et al., 2021; Vallès-Peris et al., 2021). This is the case despite their wide use and recent developments, particularly when dealing with cancer-based research (Goyal et al., 2020; Adeoye et al., 2021; Parimbelli et al., 2021). The subsequent debate has continued to gain momentum, highlighting the significant skepticism that remains among the users – and beneficiaries – of these techniques.

The extended use of big data and the biased manipulation capacities exhibited by a handful of companies have raised concerns regarding the quality of the data analyzed. The debate has reached the public sphere in the form of a variety of very direct articles warning about the manipulation faced by AI and ML techniques through small data modifications and their strong conditioning on the inherent biases that may exist, or be introduced, in the data (Metz and Smith, 2019; Aschwanden, 2020; Kaushal et al., 2020; Szabo, 2020). These warnings have swiftly moved into the ethical domain (Vellido, 2019; Gerke et al., 2020; Kerasidou et al., 2021; Lancet, 2021; Saheb et al., 2021), including data privacy considerations (Murdoch, 2021), and the epistemological one (Chin-Yee and Upshur, 2019). These shortcomings add to the actual formal challenges inherent to the application of AI and ML techniques to the analysis of different real-life settings, such as the frame problem (Yu and Kohane, 2019).

The black box quality of these techniques is one the main reasons reinforcing this distrust (Kyrimi et al., 2021; Quinn et al., 2021), with traditional statistical tests being more intuitive and already established as standard operational techniques. That is, the intuition behind survival analysis is easier to understand for physicians than that of ANN or Support Vector Machines. This is due to the mathematical complexities of the latter techniques and the fact that the resulting categorization process is not straightforward – absent specific patient profiles and an explicit description of the different factors determining model accuracy –, contrary to the more intuitive importance assigned to the explanatory variables in standard statistical regressions.

Cabitza and Campagner (2021) highlight the fact that the widespread implementation of AI and ML techniques to medical environments has taken place without generally acknowledging their potential drawbacks. These authors focus on features specific to medical datasets, particularly in terms of size and data types, that remain unaccounted for by computer scientists. More importantly, the application of ML techniques to a database should be motivated by a predetermined set of expected results following form a structured formal model. Otherwise, any categorization result derived from these techniques can be modified by randomly adding or eliminating variables while preserving a basic accuracy level.

The current paper aims at bridging – part of – the existing gap by highlighting one of the main characteristics of AI and ML techniques that allows them to outperform standard statistical analyses. While acknowledging the existence of both biases, those generated strategically and the selection bias inherent to the collection of data, AI and ML techniques preserve their categorization consistency even if the vectors of characteristics describing the evolution of patients fail to convey important information or display mistakes or mismatches in several entries. This feature constitutes an important advantage over standard statistical methods, especially when the data is organized incorrectly or sparsely. This latter problem is particularly relevant among hospitals located in developing countries, where the quality of the data retrieved is generally lower, limiting their identification and extrapolation capacities.

The quality of data, especially in developing countries, remains a concern among physicians, given the widespread existence of missing values, misprints, and mismatches (World Health Organization Regional Office for the Western Pacific, 2003 ; Lemma et al., 2020). The literature on this topic is quite extensive, highlighting the importance and consistency of the problem (Arts et al., 2002; Pezoulas et al., 2019; Koumamba et al., 2021). Thus, we offer a counterpart to the analysis of Metz and Smith (2019), emphasizing the capacity of AI and ML techniques to compensate – to a certain extent – for the existence of data misprints. We must however note that this feature is bounded by the quality of the data, since a substantial increase in the number of misprints would also limit the capacity of these techniques to remain consistent.

2.1. Supply chain environments

The structural capacity of hospital has gained relevance during the COVID-19 outbreak, which tested the limits of their ability to manage a substantial increase in the inflow of patients (Kovács and Falagara Sigala, 2021; Leite et al., 2021). As a result, the application of AI and ML techniques to manage supply chain operations within healthcare environments has increased considerably (Kumar et al., 2023; Nayeri et al., 2023a). Furthermore, the real-life decision-making processes that take place in hospitals – involving the safety of patients and their clinical evolution – have been steadily integrated within medical supply chains (Vanbrabant et al., 2023).

Hospitals face continuous coordination requirements and must deal with potential data errors in emergency situations or when unifying information across different departments (Arora et al., 2020; Rasheed et al., 2020; Vaishya et al., 2020). Errors in patient data files are common and constitute a problem consistently analyzed in the academic literature (Madden et al., 2018; Bratland et al., 2021; Millares Martin, 2022). The errors are not limited to administrative misprints but generally include medical variables (Khajouei et al., 2018; Cohen et al., 2019; Sungur et al., 2019). This problem is particularly relevant among hospitals with deficient information systems and logistic infrastructures.

Shifting to electronic patient records has been suggested as one of the main solutions to tackle this problem though it remains subject to inputting errors and has not been implemented by hospitals worldwide, those in less developed countries being more vulnerable in this respect (Ageron et al., 2018; Priestman et al., 2018; Sipanoun et al., 2022; Wowak et al., 2022). Furthermore, the main techniques applied by physicians to analyze these data are based on standard statistical methods such as survival analysis and multivariate regression.

Consequently, AI and ML techniques are being increasingly implemented in medical environments (Massie et al., 2020; Siga et al., 2020). The design of hybrid methods consisting of ML techniques and mathematical optimization models has been shown to help smooth the frictions that arise from this type of data quality constraints (Revuelta et al., 2021).

As can be intuitively understood, the capacity of AI and ML techniques to categorize patients and events correctly when dealing with evaluation frictions becomes extremely important within healthcare supply chain environments, where misclassification problems have been consistently reported (World Health Organization, 2003; Ndabarora et al., 2014; Hoxha et al., 2022). The existence of potential cumulative errors in the information gathered through the chain should therefore be incorporated into the analyses, highlighting the importance of properly formalizing sequential evaluation environments.

These cumulative frictions in information retrieval settings are tackled here from a physician-patient perspective. In particular, the current paper focuses on the categorization capacities of AI and ML techniques when dealing with actual data from clinical reports describing the main pre-transplant characteristics of patients undergoing kidney transplantation and their post-transplant evolution. Physicians focus their analyses on medical problems related to the likelihood of a successful transplantation. They do not consider supply chains explicitly but try to increase the flow of waitlisted patients into surgery while improving the expected outcomes from the transplants (Montero et al., 2021). The capacity of doctors to categorize patients according to different potential survival or graft loss scenarios improves the flow of patients through the chain and enhances the efficient use of its resources.

Supply chains in healthcare are conditioned by the potential requirements of the patients admitted into the hospital. From an information quality perspective, hospitals must be a reliable and consistent source of data. However, even with detailed data retrieved across a reasonable sample of patients, categorization problems regarding their potential evolution are common (Revuelta et al., 2021; Santos-Arteaga et al., 2021). Being able to categorize the evolution of patients would improve the management of resources throughout the chain, allowing to shift them within departments or across hospitals if needed. In this regard, the capacity of hospitals to collaborate is also generally constrained by data quality problems (Higgins, 2020; O'Halloran et al., 2020; Sauer et al., 2022).

3. Sequential online evaluation frameworks

We define now the different retrieval algorithms designed to generate the training data and the sequential online evaluation frameworks on which they are based. Consider a DM who, after performing a search online, must decide on which of the alternatives displayed by the engine to click. It has been empirically illustrated that DMs concentrate on the ten alternatives composing the first page of results (Dean, 2019) and proceed in the order provided by the engine (Lewandowski and Kammerer, 2020). That is, DMs trust the order in which the alternatives are ranked by the engine (Epstein and Robertson, 2015). They expect the initial alternatives to satisfy their preference requirements with a higher probability than those located in lower ranking positions. Thus, the probability of clicking on the link to an alternative decreases as DMs proceed through the ranking.

DMs decide whether or not to click on each alternative after reading the snippets and comparing the characteristics observed, x_i, $i = 1, 2, \ldots, 10$, with the threshold requirements defined by their subjective preferences, c_i, $i = 1, 2, \ldots, 10$. That is, DMs read the description of the first

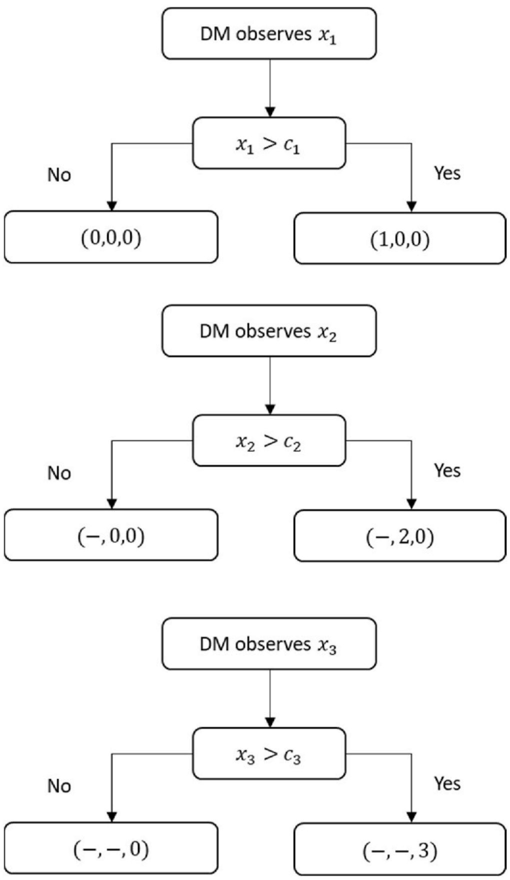


Fig. 3. Basic retrieval process composed by three alternatives.

alternative and decide whether they are interested in the potential content or prefer to continue and read the snippet of the second alternative. Clearly, if DMs are interested in the content, namely, if $x_i > c_i$, $i = 1, 2, \ldots, 10$, they click on the corresponding link. This process continues as DMs proceed through the ten alternatives, with the information assimilation capacities assumed on the DMs conditioning the results obtained.

Fig. 3 describes a basic retrieval process composed by three independent decisions made after observing the characteristics of three alternatives. The DM evaluates each alternative independently and does not incorporate the previous results obtained into his retrieval process. This feature is illustrated in the entries of the vectors that follow from the evaluation of the alternatives. If the DM clicks on the link, the resulting entry of the vector corresponds to the alternative observed. On the other hand, if the DM does not click on the link, a zero is assigned to the vector entry. The independent evaluations defining the output vectors are reported in the order observed. Thus, the entries composing the resulting matrix are given by ordered vectors where each click is registered according to the position of the alternative within the ranking. This retrieval structure describes those DMs who perform a search consisting of a given number of independent evaluations. Note that the retrieval process does not generate a complete binary decision tree.

Fig. 4 defines a complete retrieval framework based on the first three entries from a binary decision tree where a sophisticated DM acquires information while considering the previous realizations observed. In this case, each branch within the tree describes a unique path followed by the DM and determined by the number of evaluations performed and their results. In order to differentiate this retrieval process from the previous one, the evaluations defining the output vectors are reported in the initial positions, as illustrated in Fig. 4. That is, independently of the position where the alternative clicked is located, they are registered in order within the initial entries of the corresponding vector.

Consider the case where DMs set out to find ten satisfying alternatives. Five queries from the first evaluation framework – applying a threshold of ½ for all alternatives and independent uniformly distributed stochastic evaluations defined within the interval [0,1] – are presented in the 'basic' group of columns within Table 1. Note how the satisfying alternatives clicked through the retrieval process are inputted in the lower section of the vector, named 'satisfying alternatives', according to their ranking positions. Zeroes have been used to denote the absence of a click. The 'complete' group of columns describes five search queries performed within the second, more complex, retrieval framework. Note how the alternatives clicked by the DMs are grouped in the initial positions within the lower section of the retrieval vectors. In all cases, the upper section of the vectors, named 'initial evaluation', describes the

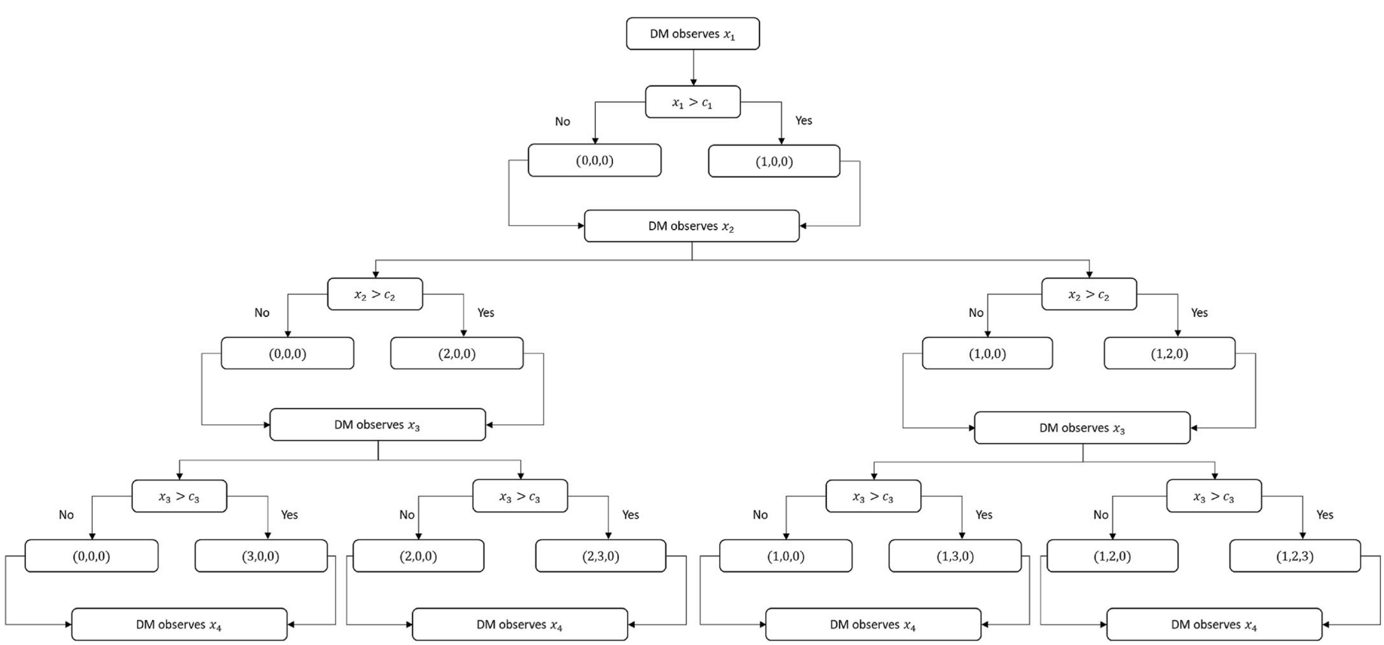


Fig. 4. Initial section of the complete retrieval process accounting for three alternatives.

Table 1
Search queries and clicking behavior across retrieval frameworks: The case with ten alternatives.

	BASIC					COMPLETE				
	DM1	DM2	DM3	DM4	DM5	DM1	DM2	DM3	DM4	DM5
Initial evaluation	0.6396	0.8457	0.4148	0.8966	0.4517	0.2947	0.4167	0.2581	0.1533	0.4730
	0.4950	0.5015	0.0945	0.8422	0.1421	0.5141	0.9439	0.4470	0.0912	0.7584
	0.9957	0.7959	0.3336	0.1360	0.7646	0.2829	0.7759	0.6406	0.8900	0.2449
	0.8853	0.1604	0.0950	0.4473	0.7478	0.2272	0.8004	0.7662	0.9290	0.3246
	0.1712	0.6523	0.0610	0.3415	0.8326	0.9431	0.1682	0.9906	0.9385	0.3925
	0.0782	0.0693	0.8874	0.9963	0.7353	0.2317	0.9422	0.6045	0.5685	0.7303
	0.6337	0.6512	0.0071	0.9843	0.1990	0.3726	0.8289	0.4000	0.8437	0.5714
	0.2241	0.5137	0.6217	0.4216	0.7227	0.5881	0.3487	0.7092	0.6195	0.9252
	0.9203	0.3310	0.5074	0.8581	0.5637	0.8932	0.8552	0.1219	0.3372	0.0100
	0.5574	0.4442	0.6187	0.9368	0.3811	0.5996	0.8991	0.9842	0.8463	0.4604
Satisfying alternatives	1	1	0	1	0	2	2	3	3	2
	0	2	0	2	0	5	3	4	4	6
	3	3	0	0	3	8	4	5	5	7
	4	0	0	0	4	9	6	6	6	8
	0	5	0	0	5	10	7	8	7	0
	0	0	6	6	6	0	9	10	8	0
	7	7	0	7	0	0	10	0	10	0
	0	8	8	0	8	0	0	0	0	0
	9	0	9	9	9	0	0	0	0	0
	10	0	10	10	0	0	0	0	0	0

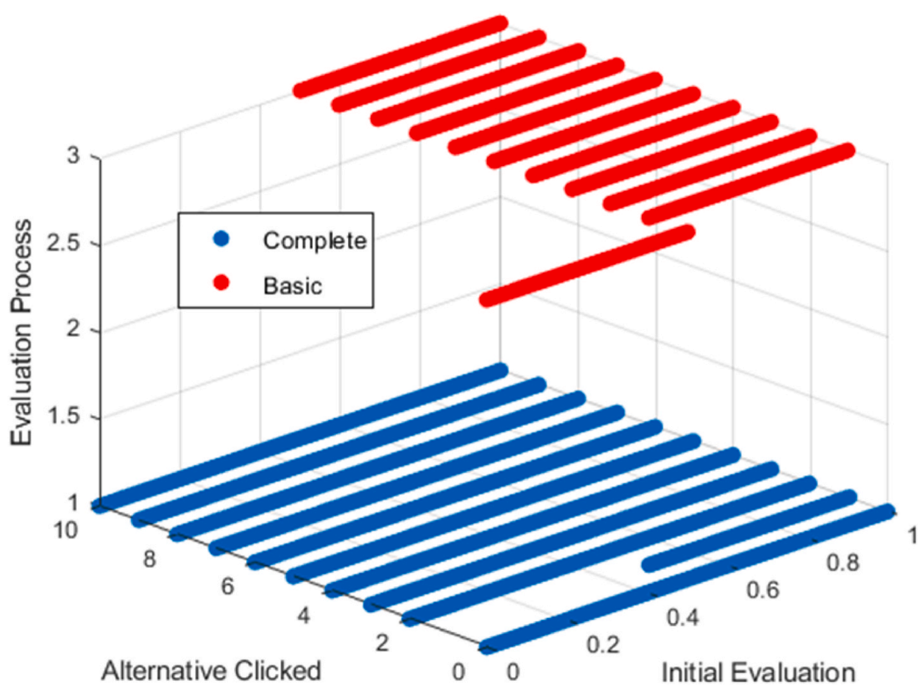


Fig. 5. Retrieval behavior derived from the basic and complete frameworks when accounting for ten alternatives.

value of the stochastic realizations that define numerically the evaluations performed by DMs.

The complexity of the algorithms formalizing the corresponding retrieval strategies differs substantially between both scenarios, as can be observed in Di Caprio et al. (2021, 2022a). In a nutshell, the basic framework consists of 10 binary decision nodes while the complete one is composed of 1023. However, a more relevant difference refers to the assimilation capacities of the DMs generating the data. For instance, in the first case, DMs cannot set out to find three satisfying alternatives within the first ten provided by a search engine. The basic retrieval structure does not incorporate this type of strategy. DMs can only evaluate the first three alternatives, or any three alternatives selected before starting the retrieval process. The second group of DMs may aim to find three satisfying alternatives out of a total of ten, since the retrieval process accounts for every potential evaluation that may be performed and the subsequent clicking behavior.

Fig. 5 provides additional intuition by describing the retrieval behavior derived from 2000 queries per framework when DMs set out to find ten alternatives satisfying their subjective preferences. The figure pairs the value of the observations with the alternatives clicked per search query when a threshold of ½ is considered per alternative. As a result, the basic framework assigns a value of zero only to those observations lower than ½.

In contrast, the complete framework cannot assign a value lower than ½ to the alternative ranked first. Note that DMs click on the first alternative whenever the initial evaluation is higher than ½. On the other hand, if DMs do not click on the first alternative, the first entry of the lower section of the retrieval vector must be different from one. The same reasoning applies to the remaining alternatives, highlighting the differences between both frameworks and illustrating how the order in

Table 2
Accuracy scores obtained from different ML techniques: 0.2 threshold scenario.

Evaluation Scenario	Features	Evaluation processes	Tree			Linear Discriminant	Quadratic Discriminant	Logistic Regression	Gaussian Naïve Bayes	Kernel Naïve Bayes	Support Vector Machine						K-nearest neighbors						Ensemble				
			Fine	Medium	Coarse						Linear	Quadratic	Cubic	Fine Gaussian	Medium Gaussian	Coarse Gaussian	Fine	Medium	Coarse	Cosine	Cubic	Weighted	Boosted Trees	Bagged Trees	Subspace Discriminant	Subspace KNN	RUSBoosted Trees
Ten	Simpl	GP/GP	51.8	51	51.6	51	49.4	51	51	50.8	51	51.7	51.8	51.7	51.8	50.8	50.5	50.9	50.6	50.9	50.9	50.8	50.7	52.2	50.8	50	51.1
		OR/GP	93	93.1	91.9	88.9	91	91	90.6	87.7	91.8	93.2	93.3	92.3	93.2	91.5	92.5	92.4	91.6	92.4	92.4	92.4	92.6	92.3	90.4	90	93
	Enhcd	GP/GP	50.1	50.9	49.8	51.2	52.3	51.2	51.4	51.1	50.3	52.6	50.7	49.5	52.5	51.3	48.9	50.5	51.1	51.6	51.1	50.8	50.8	50.9	50.6	50.1	50.8
		OR/GP	92.4	92.5	91.8	90.2	92.1	90.8	90.6	87.7	91.8	92.8	92.4	81.8	93	91.6	91.8	92.3	92.3	92.5	91.9	92.4	92	91.9	91	92.5	92.4
	Simpl	GP/OR	92.6	92.6	92.8	90.5	90.5	92.7	91.1	87.5	92.2	92.5	92.5	91.5	92.5	91.8	92.9	92.8	90.3	93	92.8	93	92.7	92.8	91.3	79.5	92.5
		OR/OR	47.6	47.2	47.8	49	46.9	49	48.8	48.6	48.3	47.6	47.7	48.2	47.3	48.3	50.1	50	49.1	49.4	50	49.2	47.1	47.9	49	50	47.2
	Enhcd	GP/OR	92.8	92.2	92.8	91.4	91.8	92.7	91.1	87.2	92.2	92.2	92.2	82.7	92.5	92.4	92	92.5	92	92.5	92.4	92.6	92.2	91.5	91.7	92.5	92.5
		OR/OR	49.4	49	49.7	48.3	48.5	48.3	48.4	47.6	48.1	48.7	49.1	48.1	46.6	48	47.9	48.9	48.2	49.9	50.1	49.2	48.8	48.8	48.4	46.7	49
Six	Simpl	GP/GP	86.8	86.8	86.8	-	-	86.7	-	87.2	87.1	86.8	86.8	86.8	87.2	86.6	87.1	86.7	86.4	86.8	86.8	87.1	87.2	63	85.8	74.9	86.8
		OR/GP	87	87	87.1	-	-	87	-	77.9	87.3	87	87	86.8	87.3	87.3	87.2	87.1	86.7	87.1	87.1	87.2	87.2	75	87.1	85	87
	Enhcd	GP/GP	87.4	86.3	86.6	-	-	87.1	-	67	86	86.7	86.9	85.4	86.2	87.2	86.6	87.3	86.4	87.1	87	87	87.3	86.9	85.8	87	86.3
		OR/GP	87.1	86.8	87.2	-	-	87.1	-	67.2	86.7	87.5	87.3	85.7	87.5	87.2	87.2	87.7	87	87.5	87.2	87.2	87	87.6	87.1	86.9	86.7
	Simpl	GP/OR	87.2	87.2	87	85.9	-	87.2	-	86.8	87.2	87.2	87.2	87.2	87.2	86.7	86.8	86.8	86.1	86.8	86.8	86.8	86.8	86.8	86	84.7	87.2
		OR/OR	86.4	86.4	86.5	86.4	-	86	-	86.8	85.9	86.3	86.4	86.1	86	86.8	86.6	86.7	82.9	86.7	86.7	86.7	85.9	86.7	86.4	74.9	86.4
	Enhcd	GP/OR	87.2	87.2	87.2	85.3	-	87.1	-	76.3	86.6	87.2	87	85.1	87.2	86.8	87.2	87.2	86.7	86.7	87.2	87.2	87.4	87.2	86.1	87.5	87.4
		OR/OR	86.6	87.1	86.9	86.3	-	87.1	-	50.3	86.1	86.4	86.5	83.9	86.8	86.8	86.5	85.9	84.3	86.5	85.8	86.1	86.9	87.1	86.5	86.2	86.7
Two	Simpl	GP/GP	68.1	68.1	68.1	-	-	68.1	-	68.1	68.1	68.1	68.1	68.1	68.1	67.4	68.1	68.1	68.1	68.1	68.1	68.1	67.3	50	67.4	58.6	68.1
		OR/GP	68.1	68.1	68.1	-	-	68.1	-	62.6	68.1	68.1	68.1	68.1	68.1	67.3	68.1	68.1	68.1	68.1	68.1	68.1	67.2	50	68.1	59.7	68.1
	Enhcd	GP/GP	68.6	67.7	67.5	-	-	67.4	-	50	67	65.8	60.2	66.8	67	67.3	68.3	68	67.8	68.2	67.8	68.7	67.9	67	67.9	68.6	67.8
		OR/GP	67	67.6	68.4	-	-	68.3	-	50	68.1	60.3	59.2	67.8	68	68.1	66.8	67.5	68	67.7	67.5	67	67.5	68.1	68.4	68	67.7
	Simpl	GP/OR	67.3	67.3	66.9	-	-	67.4	-	67.7	67.4	67.4	67.4	67.4	67.4	67.7	67.7	67.7	66.9	67.7	67.7	67.7	67.5	63.9	67.7	59.7	67.3
		OR/OR	67.2	67.2	66.4	-	-	67.2	-	67.7	67.2	67.2	67.2	67.2	67.2	67.7	67.7	67.7	66.9	67.7	67.7	67.7	67.5	63.9	67.7	54.9	67.2
	Enhcd	GP/OR	67.6	66.7	66.5	-	-	67.5	-	57.1	67	60.9	59.3	67	67	67.7	68	67.9	66.2	67.9	67.5	67.9	67.7	67.6	67.8	68.3	66.7
		OR/OR	67.9	68.2	66.9	-	-	66.3	-	50	66.8	57.5	60.4	66.6	66.5	67.7	65.4	67.3	66.2	66.4	67.2	65.7	66.9	67.5	67.7	67.3	67.2

Legend: The highest scores achieved within each evaluation scenario have been shaded.

which the data is inputted conditions the ability of the models to categorize retrieval scenarios. Clearly, threshold values close to an acceptance of 100% would generate almost identical retrieval processes, preventing AI and ML techniques from distinguishing between frameworks.

The consequences from inputting the clicks according to the ranking position of the alternative within the complete retrieval framework can be foreseen using Fig. 5. Note that, if this were the case, both frameworks would be indistinguishable when considering ten alternatives. On the other hand, they should be easier to differentiate as the number of alternatives defining the retrieval process decreases. The limits faced by standard statistical methods would however prevail, as well as the intuition following from the corresponding results.

3.1. Numerical results

We illustrate the capacity of AI and ML techniques to differentiate between both retrieval frameworks. Tables 2–4 describe the accuracy of the set of techniques applied to categorize profiles from the basic and complete frameworks based on.

- the lower section of the retrieval vectors, defining the 'simple' entries within the 'features' column, or the whole retrieval vector, defining the 'enhanced' entries within the 'features' column;
- the inputting of the clicks within the lower section of the retrieval vectors, defining the 'grouped' and 'ordered' categories within the 'evaluation processes' column – the former reports the clicks in the order observed while the latter groups them in the initial positions –;
- a variety of threshold values ranging from weaker retrieval restrictions, corresponding to a threshold of 0.2, to increasingly stricter restrictions of 0.5 and 0.8.

All tables present the results obtained when DMs set out to observe two, six and ten satisfying alternatives. We must note that a slight improvement in the categorization accuracy of the different techniques may be attained by adding the upper section of the retrieval vectors as part of the training inputs, namely, by shifting from 'simple' to 'enhanced' features. Nevertheless, the identification capacity of these techniques is mainly determined by the way the clicks are inputted in the lower section of the vector and the number of satisfying alternatives defining the retrieval process. We will therefore focus on the results obtained when comparing accuracies within the 'simple' feature category. The following analysis is based on the results obtained by the most accurate techniques within each scenario, whose entries have been shaded in the corresponding tables.

Consider the evaluation framework with two alternatives. The identification capacity of ML techniques increases as the retrieval threshold imposed by DMs shifts from 0.2 to 0.8. The intuition behind this result is clear. Within the 0.2 threshold scenario, most of the alternatives evaluated correspond to the first two composing the ranking, leaving a small margin for the DM to proceed further. That is, the probability of proceeding to the third alternative and beyond is lower compared to the other two scenarios, where the capacity of ML techniques to differentiate between DMs increases for all evaluation processes. Note that, in all scenarios, we will have a maximum of two numerical values different from zero within the corresponding retrieval vectors.

Table 3
Accuracy scores obtained from different ML techniques: 0.5 threshold scenario.

Evaluation Scenario	Features	Evaluation processes	Tree			Linear Discriminant	Quadratic Discriminant	Logistic Regression	Gaussian Naïve Bayes	Kernel Naïve Bayes	Support Vector Machine						K-nearest neighbors						Ensemble				
			Fine	Medium	Coarse						Linear	Quadratic	Cubic	Fine Gaussian	Medium Gaussian	Coarse Gaussian	Fine	Medium	Coarse	Cosine	Cubic	Weighted	Boosted Trees	Bagged Trees	Subspace Discriminant	Subspace KNN	RUSBoosted Trees
Ten	Simpl	GP/GP	49.8	50.6	50.3	51.9	-	51.9	-	50	51.9	50.9	50.4	50.1	51	50.9	48.6	50.4	49.5	50.7	50.8	49.3	51.3	51	51.4	50.1	50.6
		OR/GP	99.3	99.5	98.2	97.4	93.2	98.3	94.1	96	98.3	99.4	99.4	99.2	99.4	98.3	99.4	99.1	97.7	99.3	98.8	99.2	99.4	95.4	97	98	99.5
	Enhcd	GP/GP	49.3	50.5	50.1	50.2	-	50.2	-	50.4	51.2	50.1	51.7	49.9	51.7	51	49.1	50.6	52	50.4	50.1	50.5	50.8	49.9	50.7	49.6	50.5
		OR/GP	99	99.3	98.2	97.5	98.9	98.8	94.1	96.1	98.6	99.4	99.2	98.4	99.4	98.4	98.8	99.2	98.7	99.3	99.2	99.2	99.3	98.9	97.5	99.5	99.3
	Simpl	GP/OR	99.7	99.7	98.4	97.8	-	98.6	-	96.7	98.5	99.6	99.6	99.4	99.6	98.6	99.5	99	97.4	99.3	98.9	99.4	99.6	96.3	97.5	98	99.6
		OR/OR	51.4	50.6	51.4	49.1	50.7	49.1	49.5	50.2	48.7	51	51.2	50.8	49.8	49.3	51.2	50.3	50.5	51.2	50.3	51.8	51	51	49	50.1	50.7
	Enhcd	GP/OR	99.1	99.4	98.4	98	-	98.6	-	96.6	98.8	99.6	99.5	98	99.5	98.7	99.2	99.2	98.8	99.1	99	99.2	99.5	98.9	98	99.5	99.4
		OR/OR	49.1	48.8	50.1	49.4	49.2	49.1	49.5	49.8	49.3	50.4	50.5	52.4	51.3	50.1	51.4	51.8	51.3	51.9	52.1	53.3	49.5	50.7	49.2	51	48.7
Six	Simpl	GP/GP	95.8	95.8	94.2	-	-	88.8	-	93.3	88.8	95.7	95.6	96.1	94.6	91.8	95.7	94.3	90.5	94.6	94.3	95.7	95.8	76.1	88.2	82.5	95.8
		OR/GP	98.9	98.9	98	-	-	98.2	-	93.7	97.8	98.6	98.7	98.9	98.5	98	98.7	98.4	97.8	98.5	98.4	98.7	98.7	91.2	95.6	96.8	98.9
	Enhcd	GP/GP	98.8	98.8	98.8	-	-	98.9	-	65.1	99	98.9	99	98.5	98.9	98.7	97.3	96	93.8	96.5	96.1	96.2	50	98.9	96.6	98.2	50
		OR/GP	98.9	98.9	99	-	-	98.8	-	96.8	99	99	98.9	98.8	99	98.9	98.7	98.3	97	98.4	98.3	98.5	98.7	98.4	97.2	99	98.6
	Simpl	GP/OR	98.9	98.7	93.5	96.3	-	97.7	-	97.1	97.7	98.8	99	99.1	98.5	97.7	99	98.4	95.5	98.4	98.4	98.9	99	94.8	96.9	96.4	98.7
		OR/OR	96.3	96.3	96.1	92.8	-	96.3	-	96.3	96.1	96.3	96.3	96.3	96.3	96.3	96.3	96.2	90.9	96.3	96.3	96.3	96.3	94.8	94.5	82.3	96.3
	Enhcd	GP/OR	99	99	99	99	-	99.2	-	97.1	99.3	99.3	99.3	98.5	99.3	99.2	98.6	98.4	97.5	98.5	98.3	98.5	50	99	98.5	99.2	50
		OR/OR	99.2	99.2	99.3	97.8	-	99.3	-	96.3	99.3	99.3	99.3	98.5	99.2	98.9	97.5	96.2	94	96.2	96.2	96.2	69.5	99.1	97.4	98.9	69.6
Two	Simpl	GP/GP	87	87	87	-	-	87.1	-	87.7	87.1	87	87	87	87.4	87.2	87.4	87.4	87.4	87.4	87.4	87.4	87.3	50	87.1	62	87
		OR/GP	86.9	86.9	86.6	-	-	86.5	-	87.4	86.5	86.9	86.9	86.9	87.1	87.1	87.4	87.4	87.4	87.4	87.4	87.4	86.5	50	87.4	79.8	86.9
	Enhcd	GP/GP	87.4	86.8	86.8	-	-	87	-	49.7	86.8	86.6	86.9	87.1	86.8	87.5	88.1	86.5	87.4	87	86.6	87.8	87	87.5	87.1	87.8	86.8
		OR/GP	87.4	87.4	86.9	-	-	87	-	47.3	87.5	87.5	86.9	86.8	87.5	87.5	87.5	87.6	87.4	87.5	87.8	87.3	87.5	87.5	87.3	87.3	87.5
	Simpl	GP/OR	86	86	84.3	85.7	-	86	-	79.7	85.9	85.9	86	86	85.9	85.9	86.7	86.6	85.6	86.6	86.6	86.7	87.3	72.4	86.1	78.3	86
		OR/OR	86.3	86.3	84.2	86.1	-	86.2	-	74.2	86.1	86.3	86.3	86.3	86.3	86.6	86.6	86.6	85.7	86.6	86.6	86.6	86.1	72.4	86.5	69.7	86.3
	Enhcd	GP/OR	87.7	87.4	87	87.5	-	87.3	-	50	87.5	87.3	87.1	87	87.4	86.9	87.8	88.1	86	88.2	88.1	87.9	87.5	87.6	86.8	87.8	87.6
		OR/OR	87.3	87	86.7	86.7	-	87	-	49.9	86.6	86.7	86.8	87	86.6	86.9	87.8	87.5	86.3	87.2	87.1	87.8	87.4	86.9	86.9	87.2	86.9

Legend: The highest scores achieved within each evaluation scenario have been shaded.

With six alternatives, the highest accuracy is obtained in the 0.5 scenario. Except for the simultaneously grouped and ordered evaluations within the 0.8 scenario, the performance is clearly superior to that of the framework with two alternatives. The capacity of ML techniques to distinguish between retrieval processes is substantial in all scenarios. When considering the 0.2 case, both basic and sophisticated DMs will tend to fill the first six entries of the retrieval vector, decreasing the capacity of ML techniques to distinguish between processes. The same problem is exacerbated when considering the 0.8 scenario, with sophisticated DMs tending to proceed through the whole set of alternatives but evaluating relatively few of them among the last four. As a result, the 0.5 scenario allows for the most accurate distinction between retrieval processes.

The performance of ML techniques is similar across scenarios when considering ten alternatives. Retrieval processes cannot generally be distinguished since DMs will proceed through the ten alternatives in all scenarios. Indeed, identical evaluation processes result in a random categorization strategy leading to 50% accuracy. The main result derived from these simulations is the capacity of ML techniques to categorize DMs through the distribution of outcomes within the retrieval vectors. That is, the structure of the vector provides sufficient information to correctly identify the type of retrieval process.

Fig. 6 summarizes the main results obtained and allows for a direct comparison of the threshold evaluation scenarios categorized according to the number of alternatives considered. The intuition described above is validated together with the enhanced capacity of ML techniques to categorize both types of DMs correctly when their characteristics are inputted following different patterns within the retrieval vector.

All in all, the results obtained illustrate how ML techniques do not only consider the values defining the alternatives but also the structure of their distribution through the retrieval vectors when categorizing the corresponding processes. Thus, these techniques may allow for variations –or misprints – in some vector entrances relative to a reference structure while preserving their capacity to categorize processes correctly.

4. Medical implementation

The current section defines the medical environment analyzed, which focuses on kidney transplant patients and the information retrieved by physicians describing their evolution through the process. A team of nephrologists from the Hospital Clinic of Barcelona, Spain, has selected the main variables determining the evolution of 643 kidney transplant patients from living donors through the 2006–2019 period. The mean age of the patients composing the study equals 48,1 (±13,6) years, displaying a survival rate of 93,5%, and a death-censored graft survival of 88.6%, with a mean follow up period of 71,6 (±44,7) months. The variables selected, which extend through the three stages composing the transplantation process, are described below. The set of inputs is given by the variables defined before and during the transplant while its outcomes constitute the output set. The process incorporates both quantitative and categorical variables. Note that the binary categorical variables (0/1) assign the absence of an event to category zero while its occurrence is categorized as one. The ethics committee of the hospital has approved the corresponding study.

- **Input variables**
  - o **Pre-transplant**

Table 4
Accuracy scores obtained from different ML techniques: 0.8 threshold scenario.

Evaluation Scenario	Features	Evaluation processes	Tree			Linear Discriminant	Quadratic Discriminant	Logistic Regression	Gaussian Naïve Bayes	Kernel Naïve Bayes	Support Vector Machine						K-nearest neighbors						Ensemble				
			Fine	Medium	Coarse						Linear	Quadratic	Cubic	Fine Gaussian	Medium Gaussian	Coarse Gaussian	Fine	Medium	Coarse	Cosine	Cubic	Weighted	Boosted Trees	Bagged Trees	Subspace Discriminant	Subspace KNN	RUSBoosted Trees
Ten	Simpl	GP/GP	49.3	48.4	48.5	-	-	48.3	-	49.9	48.6	48	50.3	49.1	48.4	48.8	50.3	49.4	48.8	49.8	49.2	50.1	48.6	49	48.2	50	48.4
		OR/GP	92.9	92.9	92.8	91.6	-	91.5	-	92.8	91.4	92.6	92.9	93.1	92.7	92.6	92.9	92.7	91.2	92.5	92.7	92.8	92.9	80	91.6	79.3	92.9
	Enhcd	GP/GP	51	50	50.5	-	-	51.3	-	50	51.6	50.3	50.2	49.9	49.7	51.7	50.5	48.8	49.4	49.3	49	49.7	50.6	50.2	51	51.2	50.1
		OR/GP	93	93.2	92.8	91.5	-	91.4	-	92.7	91.8	92.9	92.8	87.5	92.7	92.3	91.8	92	90.5	92.2	90.9	92.3	92.9	89.6	91.6	93.2	93
	Simpl	GP/OR	91.2	91.7	91.5	89.5	-	89.1	-	82.2	91	92	91.7	91.8	91.9	90.4	91.4	91.3	88.8	91.2	91.3	91.4	90.9	78.1	89.5	79.5	91.7
		OR/OR	51.9	52.1	52.5	51.7	53.1	51.9	53	52	51.9	54	52.1	52.2	53.1	51.2	51.9	51.8	51.2	52	51.7	52.1	54	53.1	51.9	50.5	52.1
	Enhcd	GP/OR	91.1	91.7	91.5	89.8	-	88.7	-	82.3	91	91.4	91.4	85.3	91.5	90.2	88.9	89.8	89.3	89.9	89	90.2	91.8	89.8	89.5	91.4	91.7
		OR/OR	50.6	51.2	50.7	51.2	52.7	51.7	52.9	52.1	51	53.1	51.4	53	52.2	50.5	53.6	53.7	53.1	52.9	52.9	53.4	52.8	50.8	51.5	51.9	51.2
Six	Simpl	GP/GP	78.4	78.6	78.2	-	-	70.8	-	69.8	70.6	78.5	55.3	78.6	78.2	72.4	78.6	78.5	76.7	78	78.5	78.4	78.5	61.8	72.3	60.4	78.5
		OR/GP	92.8	92.8	92.6	-	-	90.8	-	92	92.3	92.8	92.5	92.8	92.7	91.4	92.8	92.7	89.9	92.8	92.7	92.9	92.8	65.9	90.2	83	92.8
	Enhcd	GP/GP	100	100	100	-	-	99.9	-	94.7	99.9	100	100	98.1	99.9	99.7	98.3	97.2	95.5	97.4	97.4	97.5	50	99.9	97.5	99.6	50
		OR/GP	100	100	100	-	-	99.9	-	93	99.9	100	100	98.4	99.9	99.8	99.2	98.6	96.9	98.9	98.5	99	50	99.7	98	99.8	50
	Simpl	GP/OR	93	93	81.2	88	-	87.5	-	87.5	87.8	93	92.8	93	92.2	87.8	93.1	93	90.1	93	93	93	93	77.9	87.9	79	93
		OR/OR	78.6	79	77.8	79.1	-	79.1	-	79.1	79.1	79.1	79	78.5	79	79.1	79	79	78.8	79	79	78.8	79	73.9	79.1	63	79
	Enhcd	GP/OR	100	100	100	99.3	-	99.9	-	99.9	100	100	100	97.8	100	99.9	99.8	99.3	97.1	99.1	99.3	99.5	50	99.7	99	100	50
		OR/OR	100	100	100	99	-	100	-	79.6	100	100	100	97	99.9	99.8	98.7	96.7	93.4	96.5	96.8	97.5	50	99.9	98.3	99.8	50
Two	Simpl	GP/GP	90.2	90.2	90.2	-	-	90.2	-	90.2	89.5	90.2	90.2	90.2	88.8	88.8	89.6	89.6	88	89.6	89.6	89.6	90.2	50	86.7	75.3	90.2
		OR/GP	91	91	91	-	-	91.3	-	91.2	90.8	91.2	91	91	89.6	90.1	91.2	91.2	89.5	91.2	91.2	91.2	91	50	87.2	80.1	91
	Enhcd	GP/GP	98	98	98.4	-	-	98.5	-	50	97.8	97.8	98	98.4	97.8	97.8	98.1	97.9	98	98	97.9	98	97.8	98.3	97.2	98	98.2
		OR/GP	98	97.8	97.9	-	-	97.8	-	50	97.8	97.8	97.8	97.7	97.8	97.8	97.8	97.5	97.8	97.5	97.4	97.6	97.9	97.8	97.4	98	97.6
	Simpl	GP/OR	91.8	91.8	78.8	91	-	91.9	-	86	90.7	91.7	91.7	91.7	91.7	91.9	92	92	89.8	92	92	92	91.8	73.7	90.7	77	91.8
		OR/OR	90.8	90.8	78.8	90.8	-	90.8	-	86	90.8	90.8	90.8	90.8	90.8	90.8	90.8	90.8	89.6	90.8	90.8	90.8	90.8	73.7	90.8	72.9	90.8
	Enhcd	GP/OR	98.2	98.4	98.5	98.2	-	98.6	-	47.2	98.5	98.5	98.5	98.7	98.4	98.5	98.2	98.5	98	98.6	98.5	98.1	98.3	98.3	98.3	98.3	98.3
		OR/OR	98.3	98.3	98.4	98.3	-	98.4	-	46.4	98.5	98.5	98.5	98.5	98.5	98.5	98.3	98.4	98	98.5	98.5	98.3	98.5	98.5	98.3	98.5	98.4

Legend: The highest scores achieved within each evaluation scenario have been shaded.

a) **Age at transplant (RexAge).** Potential risks derived from the transplant are higher among older patients.
b) **Compatibility type (TypeComp).** Describes the type of immunological incompatibilities existing before the transplant. The categorization of patients is based on their immunological risk:
   0 = compatible;
   1 = DS (desensitization protocol);
   2 = Exchange (cross match program applied to immunologically incompatible patients);
   3 = Exchange + DS.
c) **Type of desensitization protocol** (0, 1, 2, …,5) **(TypeDS).** Categorical variable describing the desensitization protocol applied after being triggered by a positive cross match value.
d) **Number of previous transplants (PreviousTx).** The immunological risk of patients increases with the number of transplants.
e) **ABOiPCMKPD (ABOi).** Defines the type of transplant based on its compatibility and consists of the following categories
   0 = compatible;
   1 = ABOi: different blood type;
   2 = PCM: transplanted with immunological incompatibility;
   3 = KPD: cross transplant.
f) **CKD** (0, 1, 2, …,7). Categorical variable describing the main causes leading to kidney failure.

o **Transplant**

a) **Induction.** Describes treatment with any of the following drugs during the transplant
   0 = no treatment;
   1 = basiliximab;
   2 = thymo/ATGt;
   5 = alefacept.
b) **CNI** (0/1) and **mTORi** (0/1). Drugs administered both during the transplant and for life.

- **Output variables**

o **Post-transplant**

a) **Death-censored Graft Loss** (0/1). Graft loss, excluding losses caused by the death of the patient.
b) **Death** (0/1).
   A composite binary variable is defined consisting of Death-censored Graft Loss plus Death, categorized in the following classes
   Class 1: patients either dying or losing the graft;
   Class 2: patients neither dying nor losing the graft.

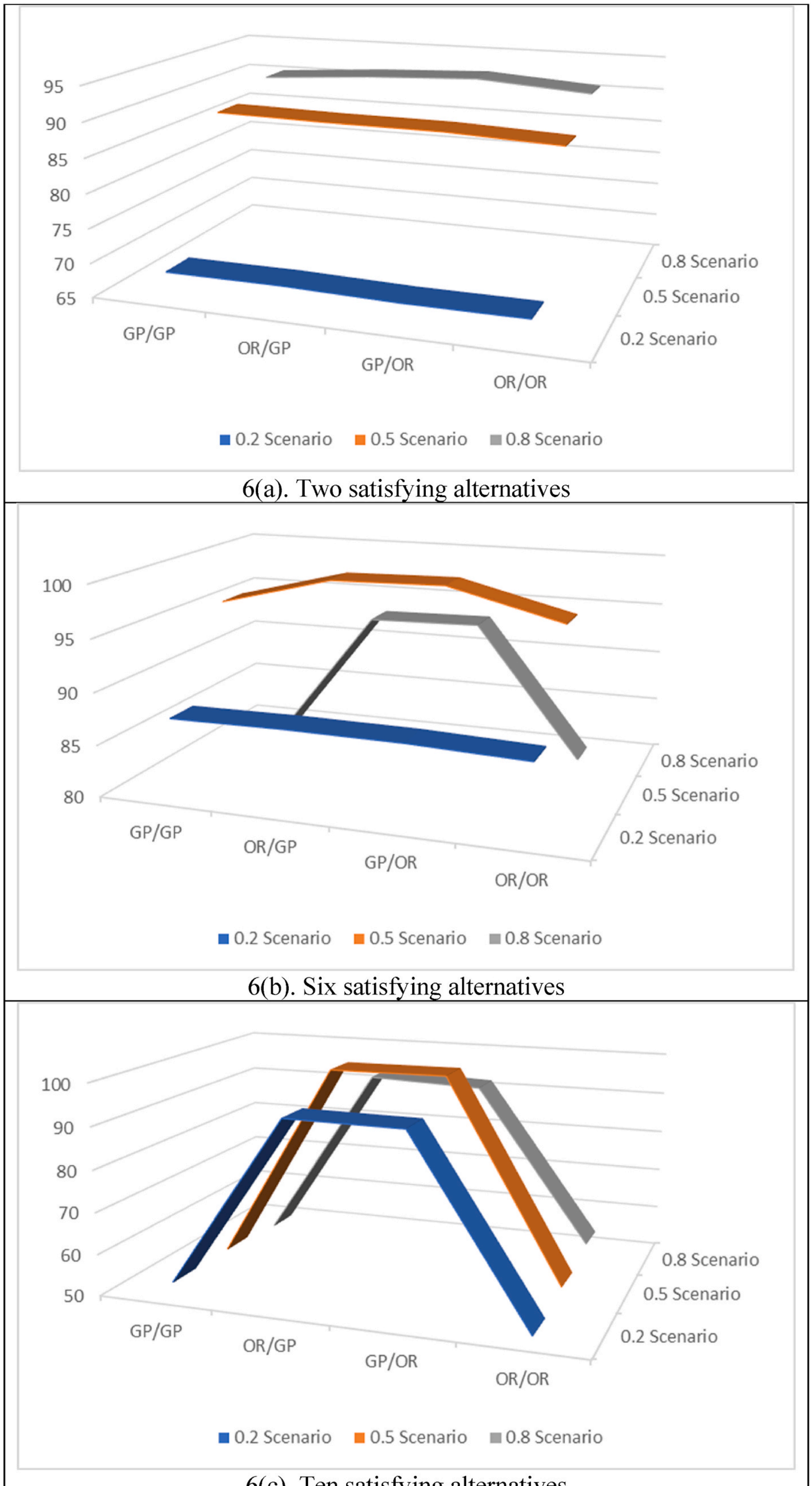
- **Modifications implemented to the observations retrieved**
  Standard scenario: considers the variables as correctly inputted by the physicians.

TypeDS	TypeComp
0	2

were reversed as follows.

TypeDS	TypeComp
2	0

Four out of these 39 patients belong to the first output class, while the other 35 belong to the second. The order of the features was preserved for any other values of TypeDS and TypeComp.



6(a). Two satisfying alternatives

6(b). Six satisfying alternatives

6(c). Ten satisfying alternatives

Fig. 6. Comparison across threshold evaluation scenarios per number of satisfying alternatives.

Table 5
Description of the entries composing the ANN confusion matrix.

		Target Class		
		1	2	
Output Class	1	True positive	False positive	Precision False Discovery Rate
	2	False negative	True negative	Negative Predictive Value False Omission Rate
		Sensitivity	Specificity	Percent of correctly classified cases
		False Negative Rate	False Positive Rate	Percent of misclassified cases

Table 6
ANN confusion matrix of the standard scenario.

		Target Class		
		1	2	
Output Class	1	10 1.6%	13 2.0%	43.5% 56.5%
	2	105 16.3%	515 80.1%	83.1% 16.9%
		8.7% 91.3%	97.5% 2.5%	81.6% 18.4%

Table 7
ANN confusion matrix of the reversed scenario.

		Target Class		
		1	2	
Output Class	1	2 0.3%	0 0.0%	100.0% 0.0%
	2	113 17.6%	528 82.1%	82.4% 17.6%
		1.7% 98.3%	100.0% 0.0%	82.4% 17.6%

Table 5 describes the entries composing the ANN confusion matrices presented in Tables 6 and 7. These latter tables summarize the identification capacities of the ANN within the standard and reversed scenarios, respectively. Note that the ANN displays the same accuracy across scenarios, as well as the same difficulties when categorizing correctly the patients belonging to the first class. Revuelta et al. (2021) highlighted this problem and proposed a hybrid Data Envelopment Analysis-ANN model to enhance the identification capacity of the neural network.

The same intuition follows from the analysis provided in Table 8, where a battery of ML techniques has been applied to both scenarios. Tables 9 and 10 present the confusion matrices corresponding to the most accurate ML techniques within the scenarios studied in Table 8. We observe similar clustering capacities exhibited by these techniques across scenarios and the same type of difficulties when identifying patients who belong to the first class.

We conclude by noting that several variables considered important by the physicians and available for a subset of patients have not been included in the analysis due to the considerable decrease in the sample size that they would impose. Moreover, the main results described remain unaffected by their inclusion. Thus, we have omitted the following variables from the pre-transplant stage: number of days in dialysis, diabetic, hypertense, and smoker patients, and reduction of immunological risk via RTX. The transplant stage does not consider the age of the donor, while the number of rejection and tumor development episodes have not been included in the post-transplant stage.

5. Confronting ML techniques and statistical analyses

We illustrate now how the identification capacities of standard statistical techniques such as a *t*-test may fail to differentiate between retrieval frameworks. To do so, we analyze the relationship between the sections of the retrieval vectors describing the alternatives clicked by the DMs. That is, we merge these sections into column vectors summarizing the alternatives clicked per framework and compute the differences in means through a paired-sample *t*-test statistic of the subsequent series.

More precisely, the vectors describing the retrieval frameworks consist of 20,000 rows, resulting from the 10 potential clicks derived from each search query and the 2000 queries simulated per framework. The null hypothesis assumes that the pairwise difference between two frameworks follows a normal distribution with zero mean and unknown variance. The statistic rejects the null hypothesis with a 5% significance whenever $h = 1$ within Table 11, which summarizes the main results obtained.

This table highlights the difficulties faced by the *t*-test statistics to tell apart both retrieval processes when considering ten alternatives. The statistic correctly identifies the vectors as being generated by different

Table 9
Description of the entries composing the ML confusion matrix.

		Predicted Class	
		1	2
True Class	1	True positive	False negative
	2	False positive	True negative

Table 10
Confusion matrices of the most accurante ML techniques across evaluation scenarios.

Standard		Reversed	
Coarse Tree		Linear Discriminant	
7	108	7	108
4	524	6	522

Table 8
Accuracy scores obtained from different ML techniques across evaluation scenarios: Medical implementation.

Evaluation Scenario	Tree			Linear Discriminant	Quadratic Discriminant	Logistic Regression	Gaussian Naïve Bayes	Kernel Naïve Bayes	Support Vector Machine						K-nearest neighbors						Ensemble				
	Fine	Medium	Coarse						Linear	Quadratic	Cubic	Fine Gaussian	Medium Gaussian	Coarse Gaussian	Fine	Medium	Coarse	Cosine	Cubic	Weighted	Boosted Trees	Bagged Trees	Subspace Discriminant	Subspace KNN	RUSBoosted Trees
Standard	76.2	80.9	82.6	81.3	79.5	81.5	78.1	82.0	82.1	81.3	79.0	81.5	82.0	82.1	72.3	80.9	82.1	81.0	81.6	77.6	80.6	81.5	82.0	48.1	62.4
Reversed	77.1	79.5	81.8	82.3	80.9	82.0	78.1	81.0	82.0	81.3	80.1	81.6	82.1	82.1	73.3	80.7	82.1	80.7	80.7	78.5	80.7	81.8	82.0	46.3	65.0

Legend: The highest scores achieved within each evaluation scenario have been shaded.

Table 11
Statistical identification of the relationship between retrieval frameworks.

Evaluation Scenario	Evaluation processes	0.2 Scenario h	0.2 Scenario p-value	0.2 Scenario t-statistic	0.5 Scenario h	0.5 Scenario p-value	0.5 Scenario t-statistic	0.8 Scenario h	0.8 Scenario p-value	0.8 Scenario t-statistic
Ten	GP/GP	0	0.5256	0.6347	0	0.9306	-0.0871	0	0.3403	0.9536
	OR/GP	0	0.9215	0.0985	0	0.4990	0.6761	1	0.0113	2.5347
	GP/OR	0	0.5367	0.6179	0	0.4389	-0.7741	0	0.1175	-1.5654
	OR/OR	0	0.7959	0.2586	0	0.9717	-0.0355	0	0.7997	0.2537
Six	GP/GP	1	0	-54.4041	1	0	-64.5065	1	0	-38.1042
	OR/GP	1	0	-74.0515	1	0	-64.7745	1	0	-32.5615
	GP/OR	1	0	-38.3412	1	0	-51.1350	1	0	-33.1133
	OR/OR	1	0	-41.8043	1	0	-54.3350	1	0	-33.2777
Two	GP/GP	1	0	-33.3370	1	0	-47.6739	1	0	-47.8506
	OR/GP	1	0	-37.6384	1	0	-50.2768	1	0	-48.0208
	GP/OR	1	0	-23.2362	1	0	-40.7329	1	0	-46.6437
	OR/OR	1	0	-23.9309	1	0	-40.7544	1	0	-46.5196

Degrees of freedom: 19,999 per scenario.

Table 12
Logistic regression: Standard scenario: Summary.

Model Summary

Step	−2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	552.491[a]	0.077	0.126

Classification Table[b]

	Observed		Predicted Output 0	Predicted Output 1	Percentage Correct
Step 1	Output	0	521	7	98.7
		1	109	6	5.2
	Overall Percentage				82.0

[a] Estimation terminated at iteration number 20 because maximum iterations has been reached. Final solution cannot be found.
[b] The cut value is .500.

Table 13
Logistic regression: Reversed scenario: Summary.

Model Summary

Step	−2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	552.491[a]	0.077	0.126

Classification Table[b]

	Observed		Predicted Output 0	Predicted Output 1	Percentage Correct
Step 1	Output	0	521	7	98.7
		1	109	6	5.2
	Overall Percentage				82.0

[a] Estimation terminated at iteration number 20 because maximum iterations has been reached. Final solution cannot be found.
[b] The cut value is .500.

retrieval frameworks when accounting for two and six alternatives. The variability exhibited within the entries of the vectors suffices to differentiate both frameworks. This is not the case when ten alternatives are considered. That is, the statistic identifies the differences between frameworks when dealing with dissimilar retrieval settings but faces considerable problems as the similarity of the processes increases.

On the other hand, ML techniques differentiate between both frameworks also when considering ten alternatives. Intuitively, these techniques account for the whole set of vector entries when categorizing the frameworks, while standard statistical models rely on the way individual realizations are inputted when defining the independent variables.

5.1. Medical implementation

In order to illustrate the identification results within a medical environment, we run two different types of regression analyses in SPSS. We do not focus on the general performance of the respective models, whose R square values are quite low, but on the modifications induced when reversing the independent input variables. That is, the main purpose of the analysis is highlighting the capacity of ML techniques to preserve a consistent categorization capacity, while standard statistical models display more significant modifications.

Consider the classification Tables 12 and 13, describing the performance of logistic regression within the standard and reversed scenarios, respectively. Note how the accuracy of this technique remains unchanged between scenarios. This feature was illustrated in Table 8, when applying logistic regression to the medical variables as one of the supervised ML techniques. MATLAB and SPSS deliver almost identical accuracies, though the latter allows for a more statistically detailed approach to demonstrate the functioning of this technique. The intuition behind this result follows from the analysis performed in Section 3, where the capacity of AI and ML techniques to identify patterns is determined by the behavior of the whole set of independent variables altogether. That is, the complete set of features, even their relative positioning, allow the corresponding techniques to identify the inherent patterns and classify the alternatives correctly.

This intuition is corroborated when considering the values shaded within the equation variables presented in Tables 14 and 15, particularly those composing the reversed scenario. As expected, the values displayed by the explanatory variables are slightly modified, though the general influence of TypeDS and TypeComp remains unchanged in both settings. That is, both these variables lack any significance on the corresponding output, with the differences in categories between them

Table 14
Logistic regression: Standard scenario: Variable analysis.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	RexAge	.023	.009	6.178	1	.013	1.023	1.005	1.041
	ABOi			1.206	3	.752			
	ABOi(1)	-18.934	40192.577	.000	1	1.000	.000	.000	.
	ABOi(2)	-20.228	40192.577	.000	1	1.000	.000	.000	.
	ABOi(3)	-.535	.572	.873	1	.350	.586	.191	1.798
	TypeDS			2.378	4	.667			
	TypeDS(1)	-.850	1.007	.712	1	.399	.428	.059	3.077
	TypeDS(2)	1.417	1.101	1.655	1	.198	4.126	.476	35.734
	TypeDS(3)	1.157	1.331	.757	1	.384	3.182	.234	43.183
	TypeDS(5)	3.149	49155.353	.000	1	1.000	23.304	.000	.
	TypeComp			.000	1	1.000			
	TypeComp(1)	19.554	40192.577	.000	1	1.000	310470347.836	.000	.
	PreviousTx	.458	.160	8.189	1	.004	1.582	1.155	2.165
	CKD			3.777	7	.805			
	CKD(1)	-.019	.474	.002	1	.968	.981	.388	2.485
	CKD(2)	.087	.449	.037	1	.847	1.090	.452	2.630
	CKD(3)	.032	.352	.008	1	.927	1.033	.518	2.059
	CKD(4)	-.400	.460	.757	1	.384	.670	.272	1.650
	CKD(5)	.351	.390	.811	1	.368	1.420	.662	3.048
	CKD(6)	.311	.559	.311	1	.577	1.365	.457	4.082
	CKD(7)	-.309	.537	.332	1	.565	.734	.256	2.103
	Induction			4.460	3	.216			
	Induction(1)	.537	.426	1.593	1	.207	1.711	.743	3.940
	Induction(2)	.422	.442	.910	1	.340	1.524	.641	3.626
	Induction(3)	2.175	1.089	3.988	1	.046	8.804	1.041	74.449
	CNI			2.261	2	.323			
	CNI(1)	-1.201	.799	2.261	1	.133	.301	.063	1.440
	CNI(2)	-20.450	16220.114	.000	1	.999	.000	.000	.
	mTORI			4.016	1	.045			
	mTORI(1)	-.694	.346	4.016	1	.045	.499	.253	.985
	Constant	-2.055	1.032	3.964	1	.046	.128		

a. Variable(s) entered on step 1: RexAge, ABOi, TypeDS, TypeComp, PreviousTx, CKD, Induction, CNI, mTORI.

arising from the shift in the values across matrix columns. Logistic regression incorporates the corresponding modifications per category into the analysis, though their general significance as well as that of the other variables remains unchanged.

The effects are substantially stronger when considering multivariable linear regression models, a statistical method commonly applied by researchers across medical disciplines (Bevan et al., 2022; Poku et al., 2022; Varady et al., 2023). In this case, the independence assumed on the explanatory variables implies that the coefficients would be severely affected when shifting values across columns. We can observe this feature when comparing the rows shaded within Tables 16 and 17, describing both the main variables directly affected and ABOi. Leaving data collinearity aside, we must highlight the larger Beta coefficient displayed by TypeDS and the negative significance exhibited by ABOi within the reversed scenario. These biases are both evident and relevant, particularly when noting that ABOi is not significant in any of the previous scenarios analyzed.

6. Discussion

We discuss in detail the main consequences derived from the results obtained within both medical and industrial settings. As is generally the case with every new technology, there is a set of potential pros and cons arising from its implementation, which are exacerbated when interacting within the medical domain (Aung et al., 2021). One of the main problems that must be tackled is the fact that even the medical community remains highly suspicious of the implementation of AI and ML techniques. Despite the reticence of physicians, these techniques remain a consistent reference tool in medical supply chain and evaluation environments. For instance, ML techniques are being currently applied to estimate the importance of different factors on the evolution of kidney transplant patients (Massie et al., 2020; Siga et al., 2020). However, there seems to remain a sense of distrust regarding the quality of these techniques and their manipulability, which makes them much less popular than standard statistical methods such as survival analysis or multivariable regression models.

Healthcare supply chains display a high degree of heterogeneity since they must deal with the stochastic inflow of patients and their potential medical conditions and evolution together with logistic and supplier restrictions common to industrial supply chains. Both types of frameworks must be simultaneously considered and analyzed. The data retrieved is also quite heterogeneous, ranging from clinical records of patients to standard supplier operations, which are largely conditioned by the state of the patients, their expected evolution, and the subsequent effect on the length of the waiting lists. These variables impose specific requirements regarding medical supplies at several points throughout the chain. Thus, when evaluating the different elements composing healthcare supply chains, the sources of information vary and are subject to errors, which would condition the outcomes derived from standard statistical methods.

The current paper has illustrated through a series of numerical simulations – based on the online information retrieval behavior of DMs – how AI and ML techniques can consistently differentiate qualities of sequential evaluation processes that remain unidentified by standard statistical methods. These results highlight the capacity of ML techniques to consider the whole structure of the retrieval vector – beyond specific values – when categorizing the search queries made by DMs. Note that we have extrapolated the intuition derived from a structured

Table 15
Logistic regression: Reversed scenario: Variable analysis.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	RexAge	.023	.009	6.178	1	.013	1.023	1.005	1.041
	ABOi			.623	3	.891			
	ABOi(1)	-20.351	40192.104	.000	1	1.000	.000	.000	.
	ABOi(2)	-20.487	40192.104	.000	1	1.000	.000	.000	.
	ABOi(3)	-.794	1.105	.517	1	.472	.452	.052	3.940
	TypeDS			.777	3	.855			
	TypeDS(1)	-.850	1.007	.712	1	.399	.428	.059	3.077
	TypeDS(2)	.260	.949	.075	1	.784	1.297	.202	8.321
	TypeDS(5)	1.731	49154.967	.000	1	1.000	5.649	.000	.
	TypeComp			1.655	2	.437			
	TypeComp(1)	20.971	40192.104	.000	1	1.000	1280865958.555	.000	.
	TypeComp(2)	1.417	1.101	1.655	1	.198	4.126	.476	35.734
	PreviousTx	.458	.160	8.189	1	.004	1.582	1.155	2.165
	CKD			3.777	7	.805			
	CKD(1)	-.019	.474	.002	1	.968	.981	.388	2.485
	CKD(2)	.087	.449	.037	1	.847	1.090	.452	2.630
	CKD(3)	.032	.352	.008	1	.927	1.033	.518	2.059
	CKD(4)	-.400	.460	.757	1	.384	.670	.272	1.650
	CKD(5)	.351	.390	.811	1	.368	1.420	.662	3.048
	CKD(6)	.311	.559	.311	1	.577	1.365	.457	4.082
	CKD(7)	-.309	.537	.332	1	.565	.734	.256	2.103
	Induction			4.460	3	.216			
	Induction(1)	.537	.426	1.593	1	.207	1.711	.743	3.940
	Induction(2)	.422	.442	.910	1	.340	1.524	.641	3.626
	Induction(3)	2.175	1.089	3.988	1	.046	8.804	1.041	74.449
	CNI			2.261	2	.323			
	CNI(1)	-1.201	.799	2.261	1	.133	.301	.063	1.440
	CNI(2)	-20.450	16220.114	.000	1	.999	.000	.000	.
	mTORI			4.016	1	.045			
	mTORI(1)	-.694	.346	4.016	1	.045	.499	.253	.985
	Constant	-2.055	1.032	3.964	1	.046	.128		

a. Variable(s) entered on step 1: RexAge, ABOi, TypeDS, TypeComp, PreviousTx, CKD, Induction, CNI, mTORI.

Table 16
Multivariable regression: Standard scenario.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
MVR	.246[a]	.061	.047	.374

a. Predictors: (Constant), mTORI, ABOi, RexAge, CNI, PreviousTx, Induction, CKD, TypeDS, TypeComp

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
MVR	(Constant)	.127	.130		.982	.326	-.127	.382
	RexAge	.002	.001	.087	2.146	.032	.000	.005
	ABOi	.010	.056	.022	.171	.864	-.100	.119
	TypeDS	.070	.024	.134	2.852	.004	.022	.118
	TypeComp	-.060	.078	-.101	-.774	.439	-.213	.093
	PreviousTx	.077	.023	.144	3.396	.001	.032	.121
	CKD	.001	.008	.003	.072	.943	-.015	.016
	Induction	.023	.022	.043	1.036	.301	-.021	.067
	CNI	-.121	.104	-.046	-1.162	.246	-.325	.083
	mTORI	-.068	.037	-.076	-1.845	.066	-.140	.004

a. Dependent Variable: Output

Table 17
Multivariable regression: Reversed scenario.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.246[a]	.061	.047	.374

a. Predictors: (Constant), mTORI, ABOi, RexAge, CNI, PreviousTx, Induction, CKD, TypeComp, TypeDS

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
MVR	(Constant)	.124	.130		.958	.338	-.130	.379
	RexAge	.002	.001	.086	2.132	.033	.000	.005
	ABOi	-.083	.032	-.189	-2.573	.010	-.146	-.020
	TypeDS	.096	.037	.208	2.576	.010	.023	.169
	TypeComp	.001	.040	.002	.030	.976	-.078	.080
	PreviousTx	.078	.023	.146	3.460	.001	.034	.122
	CKD	.001	.008	.008	.185	.853	-.014	.017
	Induction	.026	.022	.048	1.158	.247	-.018	.070
	CNI	-.122	.104	-.047	-1.173	.241	-.326	.082
	mTORI	-.072	.037	-.081	-1.950	.052	-.145	.001

a. Dependent Variable: Output

framework where the distribution of zeros and numerical values allow ML techniques to categorize retrieval processes correctly.

In this regard, modifying the value of a subset of vector entries does not affect the categorization accuracy of ML techniques. The resulting intuition has been validated using actual data retrieved from a cohort of patients undergoing kidney transplantation. That is, AI and ML techniques display a more consistent behavior than standard statistical methods in the presence of data misprints and errors. The results derived from these analyses condition the flow of patients through the medical chain – including those waitlisted –, and the subsequent requirements imposed on its operational components.

The analysis of the interactions taking place among the different components of the chain can be undertaken from a statistical viewpoint using a variety of techniques (Chin et al., 2020; Benzidia et al., 2021). Among these, structural equations constitute one of the main statistical methods applied to analyze the behavior of supply chains (Mardani et al., 2020; Belhadi et al., 2021). These approaches complement those based on standard optimization models, which have made extensive use of AI and ML techniques in healthcare-related environments (Abualigah et al., 2023; Nayeri et al., 2023b). The complementary interactions across research areas suggested also in the current paper should be particularly relevant in analytical settings with low-quality data or when integrating the databases of different hospitals or information from a variety of sources within the chain.

7. Conclusion

AI and ML techniques consider each observation as a set of predictor values ordered within a vector together with the class to which the observation belongs. We have illustrated the capacity of ANN and ML techniques to overcome the identification problems that result from the incorrect inputting of a subset of features. This quality constitutes an important advantage over standard statistical methods, particularly when the data is organized incorrectly or sparsely.

The structural framework conditioning the values derived from the simulations performed builds on a series of predetermined retrieval patterns coupled with a set of initial random realizations. The stochastic structure defining the retrieval vectors goes well beyond the simple generation of random numbers used to fill the entries of the vectors. This feature allows us to generate and evaluate a variety of retrieval profiles determined by the subjective preferences and sophistication of DMs. The analysis of the retrieval patterns generated when modifying the characteristics of DMs would link the current setting to the scenarios considered by the literature on recommender systems.

While industrial supply chain applications of the current framework may be defined, the human factor responsible for the inputting of data in hospitals together with the circulation of patient records across different sections constitute an environment prone to the emergence of misprints. That is, the interactions taking place with different components of the medical chain imply dealing with sources of information of varying quality, which would condition the subsequent results obtained. This latter problem is particularly relevant in developing countries, where the quality of the data gathered is generally lower, limiting the capacity of hospitals to exploit the results derived from its analysis.

The current results should not be interpreted as a superiority display of ML techniques relative to standard statistical analyses. Indeed, the former techniques make extensive use of standard statistical methods to improve their classification processes. The results should however invite physicians and scholars to implement and compare both types of techniques, especially in situations where data quality could be compromised. This said, it should be clear that a substantial decrease in data quality would also lead to misidentification problems within ML environments, limiting their accuracy and potential implementation. Nevertheless, when considering relatively small frictions, AI and ML techniques can be used as a validation tool to complement the main results displayed by standard statistical models, particularly in categorization settings aimed at performing basic extrapolations.

CRediT authorship contribution statement

Francisco Javier Santos Arteaga: Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Debora Di Caprio:** Writing – original draft, Methodology, Conceptualization. **Madjid Tavana:** Writing – original draft, Methodology. **David Cucchiari:** Data curation. **Josep M. Campistol:** Data curation. **Federico Oppenheimer:** Data curation. **Fritz Diekmann:** Data curation. **Ignacio**

Revuelta: Writing – original draft, Methodology, Data curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

Dr. Francisco J. Santos-Arteaga is grateful for the support received from the María Zambrano contract from the Universidad Complutense de Madrid financed by the Ministerio de Universidades with funding from the European Union Next Generation program. Research for this article was supported by funding from Mutua Madrileña Foundation (code number AP176572021). This study has also been partially funded by the Redes Tematicas De Investigacion Cooperativa En Salud (RED-INREN, RD16/0009/0023) and Redes de Investigación Cooperativa Orientadas a Resultados en Salud (RICORS2040, RD21/0005/0003), Fondo Europeo de Desarrollo Regional (FEDER) "Una manera de hacer Europa," and Secretaria d'Universitats i Recerca and CERCA Programme del Departament d'Economia i Coneixement de la Generalitat de Catalunya (2017-SGR-1331).

References

Abualigah, L., Hanandeh, E.S., Zitar, R.A., Thanh, C.L., Khatir, S., Gandomi, A.H., 2023. Revolutionizing sustainable supply chain management: a review of metaheuristics. Eng. Appl. Artif. Intell. 126, 106839 https://doi.org/10.1016/j.engappai.2023.106839.

Adeoye, J., Yan Tan, J., Choi, S.-W., Thomson, P., 2021. Prediction models applying machine learning to oral cavity cancer outcomes: a systematic review. Int. J. Med. Inf. 154, 104557 https://doi.org/10.1016/j.ijmedinf.2021.104557.

Ageron, B., Benzidia, S., Bourlakis, M., 2018. Healthcare logistics and supply chain–issues and future challenges. Supply Chain Forum Int. J. 19 (1), 1–3. https://doi.org/10.1080/16258312.2018.1433353.

Arora, N., Banerjee, A.K., Narasu, M.L., 2020. The role of artificial intelligence in tackling COVID-19. Future Virol. https://doi.org/10.2217/fvl-2020-0130.

Arts, D.G., De Keizer, N.F., Scheffer, G.J., 2002. Defining and improving data quality in medical registries: a literature review, case study, and generic framework. J. Am. Med. Inf. Assoc. 9 (6), 600–611. https://doi.org/10.1197/jamia.m1087.

Aschwanden, C., 2020. Artificial intelligence makes Bad medicine even Worse. Wired, 01.10. https://www.wired.com/story/artificial-intelligence-makes-bad-medicine-even-worse/.

Aung, Y., Wong, D., Ting, D., 2021. The promise of artificial intelligence: a review of the opportunities and challenges of artificial intelligence in healthcare. Br. Med. Bull. 139 (1), 4–15. https://doi.org/10.1093/bmb/ldab016.

Bae, S., et al., 2020. Machine learning to predict transplant outcomes: Helpful or hype? A national cohort study. Transpl. Int. https://doi.org/10.1111/tri.13695.

Bajwa, J., Munir, U., Nori, A., Williams, B., 2021. Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthcare Journal 8 (2), e188–e194. https://doi.org/10.7861/fhj.2021-0095.

Basu, S., 2018. Information search in the internet markets: Experience versus search goods. Electron. Commer. Res. Appl. 30, 25–37. https://doi.org/10.1016/j.elerap.2018.05.004.

Belhadi, A., Mani, V., Kamble, S.S., Khan, S.A., Verma, S., 2021. Artificial intelligence-driven innovation for enhancing supply chain resilience and performance under the effect of supply chain dynamism: an empirical investigation. Ann. Oper. Res. 1–26. https://doi.org/10.1007/s10479-021-03956-x.

Benzidia, S., Makaoui, N., Bentahar, O., 2021. The impact of big data analytics and artificial intelligence on green supply chain process integration and hospital environmental performance. Technol. Forecast. Soc. Change 165, 120557. https://doi.org/10.1016/j.techfore.2020.120557.

Bevan, G., Pandey, A., Griggs, S., Dalton, J.E., Zidar, D., Patel, S., Khan, S.U., Nasir, K., Rajagopalan, S., Al-Kindi, S., 2022. Neighborhood-level social vulnerability and prevalence of cardiovascular risk factors and coronary heart disease. Curr. Probl. Cardiol., 101182 https://doi.org/10.1016/j.cpcardiol.2022.101182.

Bratland, S.Z., Baste, V., Steen, K., Diaz, E., Bondevik, G.T., 2021. Physician factors associated with medical errors in Norwegian primary care emergency services. Scand. J. Prim. Health Care 39 (4), 429–437. https://doi.org/10.1080/02813432.2021.1973240.

Cabitza, F., Campagner, A., 2021. The need to separate the wheat from the chaff in medical informatics: Introducing a comprehensive checklist for the (self)-assessment of medical AI studies. Int. J. Med. Inf. 153, 104510 https://doi.org/10.1016/j.ijmedinf.2021.104510.

Chin, W., Cheah, J.-H., Liu, Y., Ting, H., Lim, X.-J., Cham, T.H., 2020. Demystifying the role of causal-predictive modeling using partial least squares structural equation modeling in information systems research. Ind. Manag. Data Syst. 120 (12), 2161–2209. https://doi.org/10.1108/IMDS-10-2019-0529.

Chin-Yee, B., Upshur, R., 2019. Three problems with big data and artificial intelligence in medicine. Perspect. Biol. Med. 62 (2), 237–256. https://doi.org/10.1353/pbm.2019.0012.

Cohen, R., Ning, S., Yan, M.T., Callum, J., 2019. Transfusion safety: the nature and outcomes of errors in patient registration. Transfus. Med. Rev. 33 (2), 78–83. https://doi.org/10.1016/j.tmrv.2018.11.004.

Dean, B., 2019. We analyzed 5 million Google search results. Here's what We Learned about Organic Click through Rate. Available at: https://backlinko.com/google-ctr-stats.

Di Caprio, D., Santos-Arteaga, F.J., 2022. Enhancing the pattern recognition capacity of machine learning techniques: the importance of feature positioning. Machine Learning with Applications 7, 100196. https://doi.org/10.1016/j.mlwa.2021.100196.

Di Caprio, D., Santos-Arteaga, F.J., 2021a. Combinatorial abilities and heuristic behavior in online search environments. Operations Research Perspectives 8, 100179. https://doi.org/10.1016/j.orp.2021.100179.

Di Caprio, D., Santos-Arteaga, F.J., 2021b. MLWA Enhancing the Pattern Recognition Capacity of Machine Learning Techniques: the Importance of Feature Positioning [Source Code]. https://doi.org/10.24433/CO.0135598.v1.

Di Caprio, D., Santos Arteaga, F.J., Tavana, M., 2022a. An information retrieval benchmarking model of satisficing and impatient users' behavior in online search environments. Expert Syst. Appl. 191, 116352 https://doi.org/10.1016/j.eswa.2021.116352.

Di Caprio, D., Santos-Arteaga, F.J., Tavana, M., 2022b. A new algorithm for modeling online search behavior and studying ranking reliability variations. Appl. Intell. 52, 7529–7549. https://doi.org/10.1007/s10489-021-02856-8.

Di Caprio, D., Santos Arteaga, F.J., Tavana, M., 2021. A new algorithm for modeling online search behavior and studying ranking reliability variations. Appl. Intell. https://doi.org/10.1007/s10489-021-02856-8.

Epstein, R., Robertson, R.E., 2015. The Search Engine Manipulation Effect (SEME) and its Possible Impact on the Outcomes of Elections, vol. 112. Proceedings of the National Academy of Sciences, pp. E4512–E4521. https://doi.org/10.1073/pnas.1419828112.

Gabrielli, M., Antonelli, M., Trovò, F., 2024. Adapting bandit algorithms for settings with sequentially available arms. Eng. Appl. Artif. Intell. 131, 107815 https://doi.org/10.1016/j.engappai.2023.107815.

Gerke, S., Minssen, T., Cohen, G., 2020. Ethical and legal challenges of artificial intelligence-driven healthcare. Artificial Intelligence in Healthcare 295–336. https://doi.org/10.1016/B978-0-12-818438-7.00012-5.

Goyal, M., Knackstedt, T., Yan, S., Hassanpour, S., 2020. Artificial intelligence-based image classification methods for diagnosis of skin cancer: challenges and opportunities. Comput. Biol. Med. 127, 104065 https://doi.org/10.1016/j.compbiomed.2020.104065.

Higgins, T.L., 2020. Not all databases are created equal. Crit. Care Med. 48 (12), 1891–1893. https://doi.org/10.1097/CCM.0000000000004636.

Hong, W., Xiong, Z., You, J., Wu, X., Xia, M.C.P.I.N., 2021. Comprehensive present-interest network for CTR prediction. Expert Syst. Appl. 168, 114469 https://doi.org/10.1016/j.eswa.2020.114469.

Hoxha, K., Hung, Y.W., Irwin, B.R., Grepin, K.A., 2022. Understanding the challenges associated with the use of data from routine health information systems in low-and middle-income countries: a systematic review. Health Inf. Manag. J. 51 (3), 135–148. https://doi.org/10.1177/1833358320928729.

Jain, K., Jindal, R., 2023. Sampling and noise filtering methods for recommender systems: a literature review. Eng. Appl. Artif. Intell. 122, 106129 https://doi.org/10.1016/j.engappai.2023.106129.

Jiang, L., Wu, Z., Xu, X., et al., 2021. Opportunities and challenges of artificial intelligence in the medical field: current application, emerging problems, and problem-solving strategies. J. Int. Med. Res. https://doi.org/10.1177/03000605211000157.

Joseph, J., Gaba, V., 2020. Organizational structure, information processing, and decision-making: a retrospective and road map for research. Acad. Manag. Ann. 14 (1), 267–302. https://doi.org/10.5465/annals.2017.0103.

Kaushal, A., Altman, R., Langlotz, C., 2020. Health care AI systems are biased. Scientific American November 17. https://www.scientificamerican.com/article/health-care-ai-systems-are-biased/.

Kelly, C.J., Karthikesalingam, A., Suleyman, M., et al., 2019. Key challenges for delivering clinical impact with artificial intelligence. BMC Med. 17, 195. https://doi.org/10.1186/s12916-019-1426-2.

Kerasidou, C., Kerasidou, A., Buscher, M., et al., 2021. Before and beyond trust: reliance in medical AI. J. Med. Ethics. https://doi.org/10.1136/medethics-2020-107095.

Khajouei, R., Abbasi, R., Mirzaee, M., 2018. Errors and causes of communication failures from hospital information systems to electronic health record: a record-review study. Int. J. Med. Inf. 119, 47–53. https://doi.org/10.1016/j.ijmedinf.2018.09.004.

Koumamba, A.P., Bisvigou, U.J., Ngoungou, E.B., et al., 2021. Health information systems in developing countries: case of African countries. BMC Med. Inf. Decis. Making 21, 232. https://doi.org/10.1186/s12911-021-01597-5.

Kovács, G., Falagara Sigala, I., 2021. Lessons learned from humanitarian logistics to manage supply chain disruptions. J. Supply Chain Manag. 57 (1), 41–49. https://doi.org/10.1111/jscm.12253.

Kumar, P., Chauhan, S., Awasthi, L.K., 2023. Artificial intelligence in healthcare: review, ethics, trust challenges & future research directions. Eng. Appl. Artif. Intell. 120, 105894 https://doi.org/10.1016/j.engappai.2023.105894.
Kyrimi, E., et al., 2021. Bayesian networks in healthcare: what is preventing their adoption? Artif. Intell. Med. 116, 102079 https://doi.org/10.1016/j.artmed.2021.102079.
Lancet. Artificial intelligence for COVID-19, 2021. Saviour or saboteur? The Lancet Digital Health. https://doi.org/10.1016/S2589-7500(20)30295-8. January 01.
Lancet, 2021. Power to the people. The Lancet Digital Health 3 (8), e462. https://doi.org/10.1016/S2589-7500(21)00147-3.
Leite, H., Lindsay, C., Kumar, M., 2021. COVID-19 outbreak: Implications on healthcare operations. The TQM Journal 33 (1), 247–256. https://doi.org/10.1108/TQM-05-2020-0111.
Lemma, S., Janson, A., Persson, L.-Å., Wickremasinghe, D., Källestål, C., 2020. Improving quality and use of routine health information system data in low- and middle-income countries: a scoping review. PLoS One 15 (10), e0239683. https://doi.org/10.1371/journal.pone.0239683.
Lewandowski, D., Kammerer, Y., 2020. Factors influencing viewing behaviour on search engine results pages: a review of eye-tracking research. Behav. Inf. Technol. https://doi.org/10.1080/0144929X.2020.1761450.
Li, H., Duan, H., Zheng, Y., Wang, Q., Wang, Y., 2020. A CTR prediction model based on user interest via attention mechanism. Appl. Intell. 50, 1192–1203. https://doi.org/10.1007/s10489-019-01571-9.
Madden, C., Lydon, S., Curran, C., Murphy, A.W., O'Connor, P., 2018. Potential value of patient record review to assess and improve patient safety in general practice: a systematic review. Eur. J. Gen. Pract. 24 (1), 192–201. https://doi.org/10.1080/13814788.2018.1491963.
Mardani, A., Kannan, D., Hooker, R.E., Ozkul, S., Alrasheedi, M., Tirkolaee, E.B., 2020. Evaluation of green and sustainable supply chain management using structural equation modelling: a systematic review of the state of the art literature and recommendations for future research. J. Clean. Prod. 249, 119383 https://doi.org/10.1016/j.jclepro.2019.119383.
Massie, A.B., et al., 2020. Identifying scenarios of benefit or harm from kidney transplantation during the COVID-19 pandemic: a stochastic simulation and machine learning study. Am. J. Transplant. https://doi.org/10.1111/ajt.16117.
Matheny, M., Thadaney Israni, S., Ahmed, M., Whicher, D. (Eds.), 2019. Artificial Intelligence in Health Care: the Hope, the Hype, the Promise, the Peril. NAM Special Publication. National Academy of Medicine, Washington, DC. Available at: https://nam.edu/artificial-intelligence-special-publication/.
Metz, C., Smith, C.S., 2019. Warnings of a Dark Side to A.I. In: Health Care, The New York Times. https://www.nytimes.com/2019/03/21/science/health-medicine-artificial-intelligence.html.
Millares Martin, P., 2022. Reducing medication errors still needs a health informatics solution in clinical practice. Health Technol. 12 (2), 523–526. https://doi.org/10.1007/s12553-022-00644-z.
Miranda, A., Miah, S.J., 2023. Designing an innovative unified contextual architecture for improving information retrieval service in healthcare organizations. Inf. Dev. 39 (2), 235–247. https://doi.org/10.1177/02666669211049492.
Montero, N., Toapanta, N., Pallares, N., Crespo, M., Diekmann, F., Guirado, L., Esteban, R., Codina, S., Melilli, E., Buxeda, A., Velis, G., et al., 2021. Deciphering transplant outcomes of expanded kidney allografts donated after controlled circulatory death in the current transplant era. A call for caution. Transpl. Int. 34 (12), 2494–2506. https://doi.org/10.1111/tri.14131.
Murdoch, B., 2021. Privacy and artificial intelligence: challenges for protecting health information in a new era. BMC Med. Ethics 22, 122. https://doi.org/10.1186/s12910-021-00687-3.
Nayeri, S., Khoei, M.A., Rouhani-Tazangi, M.R., Ghanavati Nejad, M., Rahmani, M., Tirkolaee, E.B., 2023a. A data-driven model for sustainable and resilient supplier selection and order allocation problem in a responsive supply chain: a case study of healthcare system. Eng. Appl. Artif. Intell. 124, 106511 https://doi.org/10.1016/j.engappai.2023.106511.
Nayeri, S., Sazvar, Z., Heydari, J., 2023b. Designing an IoT-enabled supply chain network considering the perspective of the Fifth Industrial Revolution: application in the medical devices industry. Eng. Appl. Artif. Intell. 122, 106113 https://doi.org/10.1016/j.engappai.2023.106113.
Ndabarora, E., Chipps, J.A., Uys, L., 2014. Systematic review of health data quality management and best practices at community and district levels in LMIC. Inf. Dev. 30 (2), 103–120. https://doi.org/10.1177/0266666913477430.
Nikzad–Khasmakhi, N., Balafar, M.A., Feizi–Derakhshi, M.R., 2019. The state-of-the-art in expert recommendation systems. Eng. Appl. Artif. Intell. 82, 126–147. https://doi.org/10.1016/j.engappai.2019.03.020.
Nwanosike, E.M., Conway, B.R., Merchant, H.A., Hasan, S.S., 2022. Potential applications and performance of machine learning techniques and algorithms in clinical practice: a systematic review. Int. J. Med. Inf. 159, 104679 https://doi.org/10.1016/j.ijmedinf.2021.104679.
O'Halloran, H.M., Kwong, K., Veldhoen, R.A., Maslove, D.M., 2020. Characterizing the patients, hospitals, and data quality of the eICU collaborative research database. Crit. Care Med. 48 (12), 1737–1743. https://doi.org/10.1097/CCM.0000000000004633.
Parimbelli, E., et al., 2021. A review of AI and Data Science support for cancer management. Artif. Intell. Med. 117, 102111 https://doi.org/10.1016/j.artmed.2021.102111.
Pezoulas, V., et al., 2019. Medical data quality assessment: on the development of an automated framework for medical data curation. Comput. Biol. Med. 107, 270–283. https://doi.org/10.1016/j.compbiomed.2019.03.001.
Poku, E., Harnan, S., Rooney, G., James, M.M.S., Hernández-Alava, M., Schaufler, T., Thokala, P., Fotheringham, J., 2022. The relationship between chronic kidney disease–associated pruritus and health-related quality of life: a systematic review. Clinical Kidney Journal 15 (3), 484–499. https://doi.org/10.1093/ckj/sfab218.
Priestman, W., Sridharan, S., Vigne, H., Collins, R., Seamer, L., Sebire, N.J., 2018. What to expect from electronic patient record system implementation: Lessons learned from published evidence. J. Innovat. Health Inf. 25 (2), 92–104. https://doi.org/10.14236/jhi.v25i2.1007.
Qin, J., Zhang, W., Wu, X., Jin, J., Fang, Y., Yu, Y., 2020. User behavior retrieval for Click-Through Rate prediction. In: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '20). Association for Computing Machinery, New York, NY, USA, pp. 2347–2356. https://doi.org/10.1145/3397271.3401440.
Quinn, T.P., Jacobs, S., Senadeera, M., Le, V., Coghlan, S., 2021. The three ghosts of medical AI: can the black-box present deliver? Artif. Intell. Med., 102158 https://doi.org/10.1016/j.artmed.2021.102158.
Rasheed, J., et al., 2020. A survey on artificial intelligence approaches in supporting frontline workers and decision makers for the COVID-19 pandemic. Chaos Soliton & Fractals. https://doi.org/10.1016/j.chaos.2020.110337.
Revuelta, I., Santos-Arteaga, F.J., Montagud-Marrahi, E., Ventura-Aguiar, P., Di Caprio, D., Cofan, F., Cucchiari, D., Torregrosa, V., Piñeiro, G.J., Esforzado, N., Bodro, M., et al., 2021. A hybrid data envelopment analysis-artificial neural network prediction model for COVID-19 severity in transplant recipients. Artif. Intell. Rev. 54 (6), 4653–4684. https://doi.org/10.1007/s10462-021-10008-0.
Richardson, J.P., Smith, C., Curtis, S., et al., 2021. Patient apprehensions about the use of artificial intelligence in healthcare. npj Digital Medicine 4, 140. https://doi.org/10.1038/s41746-021-00509-1.
Saheb, T., Saheb, T., Carpenter, D.O., 2021. Mapping research strands of ethics of artificial intelligence in healthcare: a bibliometric and content analysis. Comput. Biol. Med. 135, 104660 https://doi.org/10.1016/j.compbiomed.2021.104660.
Santos Arteaga, F.J., Di Caprio, D., Cucchiari, D., Campistol, J.M., Oppenheimer, F., Diekmann, F., Revuelta, I., 2021. Modeling patients as decision making units: evaluating the efficiency of kidney transplantation through data envelopment analysis. Health Care Manag. Sci. 24, 55–71. https://doi.org/10.1007/s10729-020-09516-2.
Sauer, C.M., Dam, T.A., Celi, L.A., Faltys, M., de la Hoz, M.A., Adhikari, L., Ziesemer, K. A., Girbes, A., Thoral, P.J., Elbers, P., 2022. Systematic review and comparison of publicly available ICU data sets-a decision guide for clinicians and data scientists. Crit. Care Med. 50 (6), e581–e588. https://doi.org/10.1097/CCM.0000000000005517.
Schmitt, J.B., Debbelt, C.A., Schneider, F.M., 2018. Too much information? Predictors of information overload in the context of online news exposure. Inf. Commun. Soc. 21 (8), 1151–1167. https://doi.org/10.1080/1369118X.2017.1305427.
Sharma, R., Gopalani, D., Meena, Y., 2023. An anatomization of research paper recommender system: Overview, approaches and challenges. Eng. Appl. Artif. Intell. 118105641 https://doi.org/10.1016/j.engappai.2022.105641.
Sharma, S., Rana, V., Kumar, V., 2021. Deep learning based semantic personalized recommendation system. International Journal of Information Management Data Insights 1 (2), 100028. https://doi.org/10.1016/j.jjimei.2021.100028.
Siga, M.M., et al., 2020. Prediction of all-cause mortality in haemodialysis patients using a Bayesian network. Nephrol. Dial. Transplant. https://doi.org/10.1093/ndt/gfz295.
Sipanoun, P., Oulton, K., Gibson, F., Wray, J., 2022. The experiences and perceptions of users of an electronic patient record system in a pediatric hospital setting: a systematic review. Int. J. Med. Inf. 160, 104691 https://doi.org/10.1016/j.ijmedinf.2022.104691.
Sungur, C., Sonğur, L., Çiçek, A.Ç., Top, M., 2019. Correction of patient medical record errors through a file control method. Health Policy and Technology 8 (4), 329–336. https://doi.org/10.1016/j.hlpt.2019.08.010.
Szabo, L., 2020. Artificial intelligence has come to medicine. Are patients being put at risk? Los Angeles Times. https://www.latimes.com/business/story/2020-01-03/artificial-intelligence-healthcare.
Utku Özmen, M., Yucel, E., 2019. Handling of online information by users: evidence from TED talks. Behav. Inf. Technol. 38 (12), 1309–1323. https://doi.org/10.1080/0144929X.2019.1584244.
Vaishya, R., Javaid, M., Khan, I.H., Haleem, A., 2020. Artificial Intelligence (AI) applications for COVID-19 pandemic. Diabetes Metab Syndr. https://doi.org/10.1016/j.dsx.2020.04.012.
Vallès-Peris, N., Barat-Auleda, O., Domènech, M., 2021. Robots in healthcare? What patients Say. Int J Environmental Research Public Health 18 (18), 9933. https://doi.org/10.3390/ijerph18189933.
Vanbrabant, L., Verdonck, L., Mertens, S., Caris, A., 2023. Improving hospital material supply chain performance by integrating decision problems: a literature review and future research directions. Comput. Ind. Eng., 109235 https://doi.org/10.1016/j.cie.2023.109235.
Varady, N.H., Pareek, A., Eckhardt, C.M., Williams III, R.J., Madjarova, S.J., Ollivier, M., Martin, R.K., Karlsson, J., Nwachukwu, B.U., 2023. Multivariable regression: Understanding one of medicine's most fundamental statistical tools. Knee Surg. Sports Traumatol. Arthrosc. 31 (1), 7–11. https://doi.org/10.1007/s00167-022-07215-9.
Vellido, A., 2019. Societal issues concerning the application of artificial intelligence in medicine. Kidney Dis. 5, 11–17. https://doi.org/10.1159/000492428.
Verma, S., Sharma, R., Deb, S., Maitra, D., 2021. Artificial intelligence in marketing: systematic review and future research direction. International Journal of Information Management Data Insights 1 (1), 100002. https://doi.org/10.1016/j.jjimei.2020.100002.

Victorelli, E.Z., Dos Reis, J.C., Hornung, H., Prado, A.B., 2020. Understanding human-data interaction: literature review and recommendations for design. Int. J. Hum. Comput. Stud. 134, 13–32. https://doi.org/10.1016/j.ijhcs.2019.09.004.

Waring, J., Lindvall, C., Umeton, R., 2020. Automated machine learning: review of the state-of-the-art and opportunities for healthcare. Artif. Intell. Med. 104 https://doi.org/10.1016/j.artmed.2020.101822.

World Health Organization Regional Office for the Western Pacific. Improving data quality: a guide for developing countries. WHO Regional Office for the Western Pacific 2003. https://apps.who.int/iris/handle/10665/206974.

Wowak, K.D., Handley, S., Kelley, K., Angst, C.M., 2022. Strategic sourcing of multicomponent software systems: the case of electronic medical records. Decis. Sci. J. 1–18. https://doi.org/10.1111/deci.12576, 00.

Wynants, L., et al., 2020. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. BMJ. https://doi.org/10.1136/bmj.m1328.

Yu, K., Kohane, I.S., 2019. Framing the challenges of artificial intelligence in medicine. BMJ Qual. Saf. 28, 238–241. https://doi.org/10.1136/bmjqs-2018-008551.

Zanganeh, M.Y., Hariri, N., 2018. The role of emotional aspects in the information retrieval from the web. Online Inf. Rev. 42 (4), 520–534. https://doi.org/10.1108/OIR-04-2016-0121.

Zhang, J., Adomavicius, G., Gupta, A., Ketter, W., 2020. Consumption and performance: Understanding longitudinal dynamics of recommender systems via an agent-based simulation framework. Inf. Syst. Res. 31, 76–101. https://doi.org/10.1287/isre.2019.0876.

Zhao, Y., Yang, L., 2024. Constrained contextual bandit algorithm for limited-budget recommendation system. Eng. Appl. Artif. Intell. 128, 107558 https://doi.org/10.1016/j.engappai.2023.107558.