

## MACHINE LEARNING ALGORITHMS IMPLEMENTATION IN THE HEALTHCARE SYSTEM AS A PROSPECTIVE AREA FOR SCIENCE, HEALTHCARE, AND BUSINESS

*Vasylevki Valerii* <https://orcid.org/0000-0001-7981-6771>

*Stepanov Ihor* <https://orcid.org/0000-0002-5844-5972>

*Koval Roman* <https://orcid.org/0000-0002-0708-6594>

*Soputnyak Mariya* <https://orcid.org/0000-0002-6888-4047>

*Liutianska Nataliia* <https://orcid.org/0000-0001-5814-5147>

*Sheyko Vladislav* <https://orcid.org/0000-0003-3063-607X>

*Stavnychyy Taras* <https://orcid.org/0000-0001-5303-0383>

*Limited liability company BIONITY, Kyiv, Ukraine*

[byblikpeper@gmail.com](mailto:byblikpeper@gmail.com)

**Relevance.** The current state of medicine is imperfect as in every other field. Some main discrete problems may be separated in diagnostics and disease management. Biomedical data operation difficulties are a serious limiting factor in solving crucial healthcare problems, represented in the statistically significant groups of diseases. Accumulation of life science data creates as possibilities as challenges to effectively utilize it in clinical practice. Machine learning-based tools are necessary for the generation of new insights and the discovery of new hidden patterns especially on big datasets. AI-based decisions may be successfully utilized for diagnosis of diseases, monitoring of general health, prediction of risks, treatment solutions, and biomedical knowledge generation.

**Objective.** To analyze the potential of machine learning algorithms in healthcare on exact existing problems and make a forecast of their development in near future.

**Method.** An analytical review of the literature on keywords from the scientometric databases Scopus, PubMed, Wiley. Search depth 7 years from 2013 to 2020.

**Results.** Analyzing the current general state of the healthcare system we separated the most relevant problems linked to diagnostics, treatment, and systemic management: diagnostics errors, delayed diagnostics (including during emergencies), overdiagnosis, bureaucracy, communication issues, and “handoff” difficulties. We examined details of the convenient decision-making process in the clinical environment in order to define exact points which may be significantly improved by AI-based decisions, among them: diagnosis of diseases, monitoring of general health, prediction of risks, treatment solutions, and biomedical knowledge generation. We defined machine learning algorithms as a prospective tool for disease diagnostics and management, as well as for new utilizable insights generation and big data processing.

**Conclusion.** Machine learning is a group of technologies that can become a cornerstone for dealing with various medical problems. But still, we have some problems to solve before the intense implementation of such tools in the healthcare system.

**Keywords:** machine learning, biomedical data science, healthcare problems, biomedical data processing.

**Relevance.** Nowadays, information technologies (IT) have a critical role in almost every field of our life, including healthcare. The development of the Biomedical Data Science sphere interconnects with gathering a significant mass of open electronic health records, developing new algorithms, and increasing computing power. Moreover, Biomedicine is one of the fastest-growing areas of knowledge over the past 30 years. As the statistics on PubMed show, in most areas of biology and medicine, the number of scientific articles doubles every few years. A large amount of data creates as possibilities as challenges, along with the development of data science allows structuring this information and, therefore, accesses it faster, using it more efficiently, generating new insights. It gives the possibility to solve a wide range of existing problems, which provides opportunities for both healthcare providers and patients. Diagnostic uncertainty is a weighty problem for healthcare providers. This phenomenon

defines as a “subjective perception of an inability to provide an accurate explanation of the patient’s health problem” [1]. According to the meta-analyzes article published in 2020, at least 0.7% of adult admissions involve a harmful diagnostic error [2]. Diagnostic uncertainty can lead to diagnostic delays, over-testing, and diagnostic errors, which can result in inadequate treatment prescription [3]. Based on the Global Health Data Exchange, 105,788 people died in 2019 from adverse effects of medical treatment, showing an increase of more than 1 percent compared to 1990. Dynamics show slow but persisting growth of iatrogenic harm for patients. We can interpret this fact as “adverse effects of healthcare system development,” which happened owing to fast data accumulation about new details and methods of diagnostics and treatment, which doctors can’t process successfully on their own in such short periods. The real numbers of iatrogenesis may be way higher than the following because to prove the adverse

effects of the treatment, full information about diagnostics and treatment should be stored and freely accessible [4]. Institute of Medicine (IOM) estimates the United States annually spends \$750 Billion in waste (approximately 30% of health care spending). IOM identified 6 waste domains: unnecessary services (\$210 billion annually); inefficient delivery of care (\$130 billion); excess administrative costs (\$190 billion); inflated prices (\$105 billion); prevention failures (\$55 billion), and fraud (\$75 billion) ([www.theatlantic.com/health/archive/2012/09/how-the-us-health-care-system-wastes-750-billion-annually/262106/](http://www.theatlantic.com/health/archive/2012/09/how-the-us-health-care-system-wastes-750-billion-annually/262106/)). Improving the quality of diagnostics and treatment by AI assistance in diagnosis of diseases, monitoring of general health, prediction of risks, treatment solutions, biomedical knowledge generation may significantly decrease costs for these domains in varying degrees. The current state of the health care system may improve with the integrated diagnostics. Integrated diagnostics is a combination of three independent diagnostic disciplines (radiology, pathology, and laboratory medicine) for therapeutic and diagnostic purposes using advanced information technology [5]. Machine learning (ML) algorithms may be an excellent tool for information collection and structurization as well as deep data analysis. Accurate diagnostics and treatment require both personal and general scientific data analysis. Personal medical data is successfully analyzed by a doctor, while the interpretation of large amounts of scientific data may significantly improve with artificial intelligence (AI). AI would also influence data storing, processing, and security as well as give some economic benefits. In this article, we will focus on the existing medical problems and their causes as well as possible solutions using ML.

## METHOD

An analytical review of the literature on keywords from the scientometric databases Scopus, PubMed, Wiley. Search depth 7 years from 2013 to 2020.

## RESULTS AND DISCUSSION

### *Relevant healthcare problems*

As in any other area of our lives, the healthcare system evolves rapidly, but its development is unaccomplished. Lately, a lot of difficulties resolve, while some of them still exist. Analyzing previous researches, we have emphasized some of the most significant problems, which can be partially solved using ML algorithms as well as other methods of biomedical data science. Among them are diagnostics errors, delayed diagnostics (including during emergencies), overdiagnosis, bureaucracy, communication issues, “handoff” difficulties, which have a significant impact on the quality of healthcare. We will describe some of them as examples of cardiovascular system diseases, neurological disorders, oncology, and kidney disease.

### *1. Cardiovascular problems*

According to the World Health Organization (WHO), cardiovascular diseases (CVDs) are the number 1 cause of death globally. Heart attacks and strokes cause four out of five cardiovascular disease deaths. The proportion of deaths from non-communicable diseases is 41 million annually, leading to 71 % of deaths, respectively. Among them, cardiovascular disease is a global problem, killing 17.9 million people annually, or 31 % of all deaths each year [6]. Early diagnosis of atherosclerosis and blood clots may prevent complications such as stroke, heart attack, or pulmonary embolism. Late detection of atherosclerotic plaques can also lead to the development of aneurysms and problems with coronary, carotid, peripheral, renal arteries [7]. Early diagnostics has a significant effect on survival the earlier it carries out, the higher the chance of recovery [8]. Thus, monitoring of biomarkers as well as other predictors of CVDs that indicate the development of blood clots and atherosclerotic plaques can significantly improve the situation and reduce the total number of deaths related to the cardiovascular system. The rate of cholesterol testing two or more times in 3 years has been growing exponentially for the last twenty years. This tendency has a positive correlation with patients’ data accumulation, which gives more possibilities for efficient monitoring of lab test dynamics. Such an approach can be useful for any chronic disease. To illustrate that cholesterol level monitoring may be helpful as a predictor of cardiovascular events. As a result, we can enhance prophylactics and prescribe a preventive treatment, such as lipid-lowering drugs, before consequences develop [9].

According to WHO, CVDs, and two other groups of diseases (cancer and infectious diseases) are the main categories of average diagnostic errors in primary health care. Errors may occur when minor warning symptoms are missed or ignored in primary care [10]. Cross-analysis of a large population sample estimate from 15,000 to 165,000 misdiagnosed cerebrovascular events annually in United States emergency departments (EDs), disproportionately representing headache or dizziness. Some cerebrovascular diseases do not diagnose immediately, which can lead to mortality or disability of the patient [11]. Approximately 9% of cerebrovascular diseases go unnoticed at the initial ED presentation. The risk of misdiagnosis is higher if the patient’s complaints are minor, and the symptoms are mild, non-specific, or transient [12]. Another important one is deep vein thrombosis (DVT). There is a high risk of missing the diagnosis in a patient with deep vein thrombosis. According to Yuhong Zhang, the missed diagnosis of DVT in the lower extremities using ultrasound is about 50% in patients without DVT symptoms [13]. The most significant complication of DVT is pulmonary embolism (PE), which is a very life-threatening condition. PE may most likely lead to death without proper management, and its diagnostics may often delay clinical practice [14].

Also, there are a lot of errors related to cardiovascular medications. The ED and acute hospital is the most common locations at high risk for medication errors [15]. According to a study in 2016, cardiovascular drugs associate with 24.7% of medical errors. Among them, leading anticoagulants - 11.3% of the errors [16]. Besides, the use of incorrect doses of the drug and unnecessary drugs for the treatment of cardiovascular diseases can lead to the development of thrombosis [17].

## 2. Neurological disorders

**2.1 Parkinson's disease.** According to the statistics, about 1% of people over age 60 have Parkinson's disease (PD), and this percentage rises with aging [18]. In the case of early diagnosis, the efficiency of pharmacological management increases, and non-pharmacological management is also possible. The combination of these two approaches helps to manage present symptoms and prolongs an active and healthy life [19]. Early diagnostics may conduct with the help of neurochemical biomarkers, such as orexin, Dopamine, Dopamine receptors, and Dopamine Transporter Activity,  $\alpha$ -Synuclein, and others [20]. PD misdiagnosing occurs in about 30% of all cases [21]. An accurate diagnosis of PD is essential both for patients care and researches associated with epidemiology, genetics, medical imaging, neurochemical biomarkers, and symptomatic and disease-modifying treatments [22].

**2.2 Multiple sclerosis.** Multiple sclerosis (MS) is a potentially disabling autoimmune disease without efficient treatment and exact etiology. It is characterized by immune-mediated attacks on the central nervous system (CNS) and following demyelination and reversible or relapsing neurological symptoms [23]. Hence, it is obviously expedient to generate new insights on early laboratory diagnostics of MS by analyzing big patient data. According to the statistics, MS affects approximately 900,000 people in the United States and 2.5 million people worldwide [24]. Following the statistics from studying in the UK, peak incidence occurred between ages 40 and 50 years and maximum prevalence between ages 55 and 60 years [25]. Early diagnosis of MS is possible as a combination of symptoms (lasting at least 24 hours) and clinical tests, including early biomarkers: oligoclonal bands, anti-MOG antibodies, antinuclear antibodies [26]. It helps to decrease the possibility of disability and to lower the secondary relapse rate [27]. Misdiagnosis of MS brings certain risks associated with not receiving early-stage treatment [28]. The diagnostic error usually occurs when different disorders that aren't associated with demyelination and inflammation processes expose symptoms typical for MS [29]. Alternative conditions may frequently suggest by the presence of "red flags" in the clinical presentation. These are atypical for MS signs, symptoms, or findings that should be detected and investigated by radiographic, clinical, or laboratory methods to reduce the possibility of MS misdiagnosis [29].

**2.3 Alzheimer's disease.** Alzheimer's disease (AD) is

a detrimental worldwide social problem. The prognosis estimates that the number of AD patients only in the USA will rise from 5 million to 14 million by 2050 [30]. According to WHO, the total number of patients with dementia may reach 82 million in 2030 and 152 million in 2050 [31]. AD should diagnose in the preclinical phase or while AD-induced mild cognitive impairment (MCI) to decrease the possibility of irreparable brain damage. Some people with MCI have returned normal cognition without dementia related to AD due to on-time diagnostics and treatment [30]. Diagnostics of AD with the help of biomarkers measurement may prevent a significant number of false-positive diagnoses, as opposed to alone guidelines diagnostics [32]. In perspective, early biomarkers will become a required part in monitoring the effects of AD treatment [30]. Detection and measurement of biomarkers are also a significant part of AD drug development. It allows identifying better compatibility between designed drug and patient for clinical trials [33]. That's why the biomarkers monitoring process should simplify in all stages: from diagnostics to treatment. Patients and people associated with diagnostics and treatment of AD should be able to get fast access to correct and relevant measured biomarkers.

## 3. Oncology

**3.1. Thyroid cancer.** Thyroid cancer (TC) is the most common endocrine cancer [34]. TC's are often overdiagnosed in the USA and South Korea, mainly by ultrasonography [35][36]. The most common way of thyroid tumor management is a radical thyroidectomy. It provokes an increase in hypoparathyroidism incidence in South Korea [36]. In 2015, the Korean Committee for National Cancer Screening Guidelines issued a recommendation against thyroid cancer screening with ultrasonography for healthy individuals [37]. Thyroid cancer has an estimated 5-year survival of 98.1% overall: 99.9% for localized disease and 55.5% for distant disease [34]. It may indicate that thyroidectomy may be irrational in some cases. So, harm from disease management may be more significant than one from the tumor itself. Fine-needle aspiration is the most common method in the diagnosis of TC. When performed, 70% of all thyroid tumors classify as benign, 4.0% as malignant, and 10% as suspicious or indeterminate, and 17% demonstrate an insufficient sample [38]. There are some non-specific biomarkers that may indicate a presence of the thyroid tumor [39] that all together and in combination with other existing diagnostic methods can more accurately suggest a possibility of TC. This approach will provide a more accurate prescription of thyroidectomy.

**3.2. Prostate Cancer.** Prostate Cancer (PC) is the second most common cancer in men. In the USA, 33,330 deaths occur from prostate cancer [34]. The implementation of the prostate-specific antigen (PSA) test increased the level of prostate cancer detection, resulting in overdiagnosis and overtreatment [40]. Undergoing radical pros-

tactomy or radiation therapy may lead to some complications (urinary symptoms, operative mortality) as well as long-term sequelae (urinary incontinence, impotence, and bowel dysfunction) [41]. In our opinion, a PSA blood test alone is not enough for the diagnosis, so we need more specific biomarkers (BM's) [42].

#### 4. Kidney disease

Chronic kidney disease (CKD) is a significant disorder that affects a lot of people around the world. Over \$1 trillion is spent globally on end-stage renal disease care [43]. Unfortunately, it's often recognizable only by laboratory abnormalities in the latest stages. Late diagnostic caused no effective kidney disease treatment development [44] and may limit the number of BM's for early disease detection. Measuring glomerular filtration rate (GFR) is a "gold standard" for CKD assessment, but not specific enough, especially during the early stages of the disease [45]. There is a bid amount of BM's associated with kidney damage and, or loss of function, which can implement via ML methods for CKD management [46]. Factors such as age, gender, race, and family history are crucial for CKD. Moreover, hypertension, smoking, diabetes mellitus, and obesity can also lead to kidney disease. It's critical for patients and doctors to notion all aspects of effective diagnostics and treatment. Normal renal senescence and physiological loss of GFR should be noticed and differentiated from life-threatening signs of CKD. Complete analysis of the patient's data: including his age, comorbidity, in complex with albuminuria, GFR, and early biomarkers of kidney damage is a potentially efficient tool in CKD diagnostics and management [47].

#### Healthcare problems and data processing

Such tools as electronic health records gave a push for biomedical data science development but still, it doesn't look like we can use that vast amounts of data fluently. According to The Joint Commission Center for Transforming Healthcare Hand-off Communications Project, "hand-off is a transfer and acceptance of patient care responsibility achieved through effective communication". The hand-off is a process of medical information transmission from one health care provider to another for treatment or diagnostic propose. There are more than 4,000 hand-offs in a typical teaching hospital every day (<https://psnet.ahrq.gov/web-mm/triple-handoff>). A lot of essential information may get lost during a patient's data transfer between healthcare providers. The electronic health record system became a grandiose invention, which solved this problem to a large extent, but it still partially exists. Another serious difficulty is receiving essential medical data for urgent patient management when his person is unidentified. As we can see in modern devices, it may be partially solved by data storage on smartphones or other devices and proper linking these local systems to the electronic health record system. It also would be great to give patients the possibility to

participate in their biomedical data replenishment, but only under the doctor's control.

Such tools as electronic health records gave a push for biomedical data science development but still, it doesn't look like we can use that vast amounts of data fluently. We mentioned many BM's that may successfully use for the improvement activities in particular diseases diagnostics and management. But to do so, a lot of scientific information should be identified, analyzed, and verified in detail before the implementation of them in clinical practice. With the help of AI, we can gather a large amount of data about potential BM's from science resources like PubMed for diagnostic improvement. ML algorithms can become an excellent instrument for such BM's significance assessment and defining their role in disease diagnostics and management. Utilizing such an approach can potentially solve problems of overdiagnosis we have shown in part about oncology, help predict and detect chronic diseases in early stages and, generally, partially solve the problems we have listed earlier.

Comprehensive analysis of the patient's data: including his age, sex, race, comorbidity in complex with utilizing suitable diagnostic methods, as well as general scientific data is a promising tool for the healthcare future.

#### Machine learning algorithms and biomedical data processing

Decision-making in medicine is a responsible and complex task that requires taking into account a huge number of factors. Depending on the field of medicine, these factors and their number may vary, but even so, we can identify the most fundamental among them:

- Patient Laboratory Data / Clinical History / Genomics Data
- Psychological state / Human conditions
- Consumption of pharmaceuticals

Modern science is trying to describe and digitize these factors. It should be noted that the assessment of such factors requires the adoption of both general information and personal data. Since a human body is a complex object, it is difficult to make deductive conclusions about its nature, therefore, personal information is valuable. On the other hand, it is rather difficult to interpret personal data, therefore, for making an informed decision, the best strategy is to focus both on patient data and global information.

In the current technological situation, artificial intelligence is not able to take over decision-making, but modern machine learning algorithms can be an excellent tool for medical professionals. Such systems should have the following properties (table1):

- Good accuracy: machine learning algorithms should have near or higher accuracy than physicians.
- Good explanation ability: output result of a system should be interpretable for physicians.
- Ability to work with missing and noisy data: it is a widespread situation in medicine when data are



missing on noisy, so algorithms should be less sensitive as possible for this situation.

- Ability to use different input data with is count minimization: it very time and cost expensive to use a lot of input variables, algorithms should be able to work with different data, whit minimal volume.

- High performance: algorithms must be able to train fast and efficiently.

- Large coverage of output variables: the more output solutions a system has to offer, the higher its value.

- High differentiation power: there are a lot of subtypes of some diseases, so it is important to create very detailed systems for their classification.

We can highlight the following areas of medicine in which machine learning is actively implemented: diagnosis of diseases, monitoring of general health, prediction of risks, treatment solutions, biomedical knowledge generation. In general, the proliferation of machine learning in clinical practice is lagging behind the potential possible, given the opportunities already available.

Diagnosis of diseases is a classical topic, which people associate with using machine learning in medicine. Formally, this task can be described as predicting a certain class from the entire class pool based on input data, in the form of attributes for each patient. Lets, patient  $P = \{p_1, p_2, \dots, p_k\}$ ,  $|P|=k$ ,  $A$  is a set of attributes,  $A = \{a_1, a_2, \dots, a_n\}$ ,  $|A|=n$ . A set of patients and their attributes form a matrix MAP, that is constructed by  $P \times A$ .

Set of classes  $C = \{c_1, c_2, \dots, c_m\}$ , for each row in MAP corresponds to some class, in the case of disease classification, each class is a specific disease. The task is based on a known predict class for new patient data, based on its attributes. The most popular classifiers are Support Vector Machine (SVM), Naive Bayes Classifiers, Random Forest, K-Nearest-Neighbour (KNN), neural network classifiers. In Table 1, we have shortly described them, their advantages and disadvantages.

A very prospective field is using external knowledge graphs for improving medical diagnosis results [48]. Knowledge graph  $g = \{N, R\}$  is a set of medical entities as nodes (N) and relations between these entities {R}.

Sometimes we do not have a set of classes, but we want to subdivide and categorize sets of attributes. Examples of clustering algorithms are K-means Clustering, Agglomerative Clustering, Multikernel Learning algorithms, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), Ordering Points To Identify the Clustering Structure (OPTICS).

**Biomedical data security**

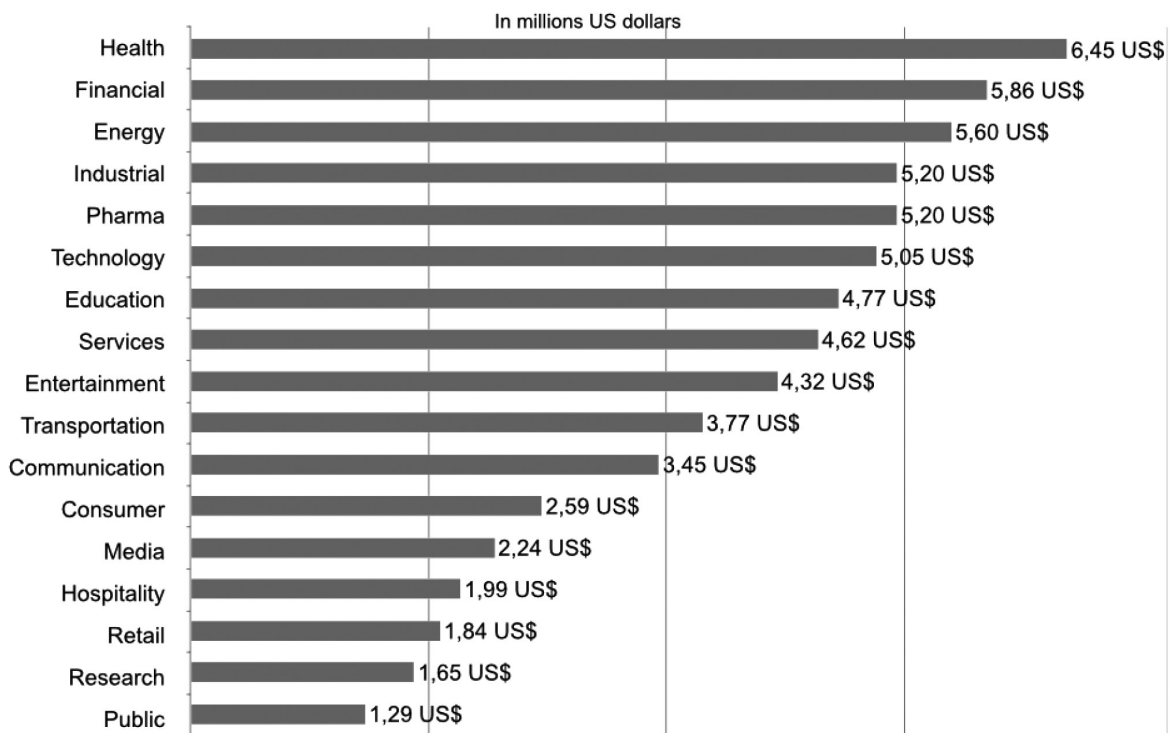
Accumulation of patients’ personal medical information, creation of new services for its archivation and processing caused the manifestation of an essential problem with data security. According to the Digital Guardian analysis published in 2019, the average total cost of a data breach by industry was the highest in the healthcare

Table 1

**The most popular machine learning technologies**

Method	Principle	Advantage	Disadvantage
Support Vector Machine [50]	Input vectors are non-linearly mapped to a very high-dimension feature space, where a linear decision surface is constructed.	<ul style="list-style-type: none"> <li>· High accuracy</li> <li>· Work well with high dimensional space</li> <li>· Use less memory</li> </ul>	<ul style="list-style-type: none"> <li>· High training time</li> <li>· Don't work well with overlapping classes</li> </ul>
Simple Bayesian classifiers [51]	Probabilistic approaches, that make assumptions about how the data is generated, and posit a probabilistic model that embodies these assumptions; then they use a collection of labeled training examples to estimate the parameters of the generative model, classification of new examples is performed with Bayes' rule.	<ul style="list-style-type: none"> <li>· Easy to implement</li> <li>· High performance</li> <li>· High explanation ability</li> </ul>	<ul style="list-style-type: none"> <li>· Work bad, when there is a dependence between variables</li> </ul>
Random forest	Random generation of a set of decision trees from a selected subset of a training set and making a conclusion about classification based on average result among them	<ul style="list-style-type: none"> <li>· Fast to train with test data</li> <li>· Good explanation ability</li> <li>· Prevents overfitting of data</li> </ul>	<ul style="list-style-type: none"> <li>· Slow in creating predictions once a model is made.</li> <li>· Must beware of outliers and holes in the data</li> </ul>
K-Nearest-Neighbour	Classification occurs by determining the class to the k nearest neighbors, determined by a certain distance function, perhaps all Euclidean distance.	<ul style="list-style-type: none"> <li>· Easy to implement</li> <li>· No training operations</li> </ul>	<ul style="list-style-type: none"> <li>· High classification time</li> <li>· Missing values and outliers sensitivity</li> </ul>
Neural Networks	Changing weights between layers with neurons in such a way as to minimize the selected loss function.	<ul style="list-style-type: none"> <li>· Good accuracy</li> <li>· Ability take into account complex relationships</li> </ul>	<ul style="list-style-type: none"> <li>· Low explanation ability</li> <li>· Possibility of local minimum</li> <li>· Possibility of overfitting</li> </ul>

Figure 1:  
Average total price of breached data in different industries



<https://digitalguardian.com/blog/whats-cost-data-breach-2019>

Fig. 1. Average total price of breached data in different industries

industry (\$429 per healthcare record) (Fig. 1) (<https://digitalguardian.com/blog/whats-cost-data-breach-2019>).

The reason for this phenomenon is that personal medical information cannot be changed once it was stolen, on the opposite of credit card or social insurance data. It can be proved with the annual report of the IBM company and Ponemon Institute, which states that the average data leakage cost in 2019 was 3,92 million dollars, while for healthcare it was 6,45 million dollars, two times higher than in any other field. From 2005 to 2019, 249,09 million people suffered from a data breach in healthcare and we can notice that the number of HACKs was consistently growing during this period [49].

Data breach in the insurance company Premera Blue Cross can be a good example of hacking. In May 2014, an employee of this company received a phishing email with a link to the document, which contained malware that allowed a malefactor to enter an internal network of the company and theft the medical data of 11 million people (<https://cutt.ly/6zDBjZG>). It's also important to protect portative devices from the access of the third person as happened in Chicago in 2013 when 4 unencrypted computers were stolen from Advocate Medical Group. 4 million patients suffering from this incident and financial loss was assessed to be billions of dollars ([www.healthcareitnews.com/news/Advocate-Health-slapped-with-lawsuit-after-massive-data-breach](http://www.healthcareitnews.com/news/Advocate-Health-slapped-with-lawsuit-after-massive-data-breach)). It should be noticed that data breaches can be caused not only by external attacks but also from the inside.

It's crucially important to constantly upgrade systems that work with biomedical data because their malfunction can lead to wrong treatment and lethal outcomes, as happened in the hospital in Germany in 2020 ([www.wired.com/story/a-patient-dies-after-a-ransomware-attack-hits-a-hospital/#:~:text=A woman seeking emergency treatment, was widely reported on Thursday](http://www.wired.com/story/a-patient-dies-after-a-ransomware-attack-hits-a-hospital/#:~:text=A woman seeking emergency treatment, was widely reported on Thursday)). AI-based systems, which we review in this article as a prospective accessory diagnostic instrument, require supporting mechanisms to confirm the authenticity of the data they operate. AI implementation in the healthcare system requires standardization of data sets it works with to prevent inaccurate results. Such systems, especially their data sets, must be protected from all types of data breaches. It's important not only to provide technical innovations in data security but also to check on the employees that have access to this information, implement the multi-layer revision of all the changes. Both for just storing patients' information and operating on it using AI-based systems, we can list some general recommendations for biomedical data security: use of anti-malware solutions and protection networks with effective firewalls; use of multi-factor authentication; use of security patch management, anti-social engineering, and phishing programs; investing in cybersecurity insurance; data encryption; creating reliable backups; investing in employee training; conducting frequent audits of the cybersecurity system.

**Contributors.** All the authors did the literature search, read, and agreed to the final manuscript.

**Declaration of interests.** We declare no competing interests.

## CONCLUSION

In this article, we have detected some severe healthcare system problems and reviewed their possible solutions using ML algorithms. Based on our analysis, the healthcare system was upgraded significantly by inventions like HER's, but still, we have target points for its improvement. Machine learning is a group of technologies that can become a cornerstone for dealing with various medical problems. Utilizing various types of AI will be useful for the understanding of the risk factors, behavioral patterns, and features of the therapeutic pathway to provide adequate in-time treatment. To implement this technique, healthcare should focus on big data accumulation and structurization, making the substrate for further investigations and market development. It's necessary to provide an interdisciplinary approach for universal AI formation, concentrating on the realization of technologies for global scientific data mining and processing. This strategy will give us the possibility to use AI as a full-fledged accessory diagnostic instrument and improve many medical issues. The application of ML algorithms, in perspective, can provide a tool for comparison outputs of laboratory and clinical studies with existing healthcare standards in order to help in the development of more advanced diagnostic methods and personalized treatment by generating new insights and detecting hidden corrections. All of the mechanisms described above will allow gathering large amounts of big data that will set a stage for the future development of biomedical data science. The development of AI-based systems requires control over the data it works with to provide accurate and reliable results. Weighty economic benefits to be expected for patients and the healthcare system in general. We believe that in the near future AI will become a fully functional diagnostic instrument, collaborating with physicians in order to provide the best quality of medical services. We believe, it will be possible for future AI's to substantially optimize healthcare reducing, making the lives of both patients and healthcare providers a bit better.

## REFERENCES

- Bhise V, Rajan SS, Sittig DF, Morgan RO, Chaudhary P, Singh H. Defining and Measuring Diagnostic Uncertainty in Medicine: A Systematic Review. *J Gen Intern Med.* 2018 Jan;33(1):103-5. DOI: 10.1007/s11606-017-4164-1.  
View at: Scopus: <https://link.springer.com/article/10.1007/s11606-017-4164-1>  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/28936618/>
- Gunderson CG, Bilan VP, Holleck JL, Nickerson P, Cherry BM, Chui P, Bastian LA, Grimshaw AA, Rodwin BA. Prevalence of harmful diagnostic errors in hospitalised adults: a systematic review and meta-analysis. *BMJ Qual Saf.* 2020 Dec;29(12):1008-18. DOI: 10.1136/bmjqs-2019-010822.  
View at: Publisher Site: <https://qualitysafety.bmj.com/content/29/12/1008>  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/32269070/>
- Singh H, Schiff GD, Graber ML, Onakpoya I, Thompson MJ. The global burden of diagnostic errors in primary care. *BMJ Qual Saf.* 2017 Jun;26(6):484-94. DOI: 10.1136/bmjqs-2016-005401.  
View at: Publisher Site: <https://qualitysafety.bmj.com/content/26/6/484>  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/27530239/>  
PubMed Central: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5502242/>
- Carver N, Gupta V, Hipskind JE. Medical Error. 2020 Oct 5. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2020 Jan. PMID: 28613514. Bookshelf ID: NBK430763  
View at: PubMed: <https://pubmed.ncbi.nlm.nih.gov/28613514/>  
PubMed Central: <https://www.ncbi.nlm.nih.gov/books/NBK430763/>
- Lippi G, Plebani M. Integrated diagnostics: the future of laboratory medicine? *Biochem Med (Zagreb).* 2020 Feb 15;30(1):010501. DOI: 10.11613/BM.2020.010501.  
View at: Publisher Site: <https://www.biochemia-medica.com/en/journal/30/1/10.11613/BM.2020.010501>  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/31839719/>  
PubMed Central: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6904966/>
- World Health Organization. Assessing National Capacity for the Prevention and Control of Noncommunicable Diseases Global Survey. *Global Survey 2015.* Geneva, 2016  
View at: Publisher Site: <https://apps.who.int/iris/handle/10665/246223>
- Fok PW, Lanzer P. Media sclerosis drives and localizes atherosclerosis in peripheral arteries. *PLoS One.* 2018 Oct 26;13(10):e0205599. DOI: 10.1371/journal.pone.0205599  
View at: Publisher Site: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0205599>  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/30365531/>  
PubMed Central: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6203409/>
- Taylor CJ, Ordóñez-Mena JM, Roalfe AK, Lay-Flurrie S, Jones NR, Marshall T, Hobbs FDR. Trends in survival after a diagnosis of heart failure in the United Kingdom 2000-2017: population based cohort study. *BMJ.* 2019 Feb 13;364:l223. DOI: 10.1136/bmj.l223.

- View at: Publisher Site: <https://www.bmj.com/content/364/bmj.l223>  
 PubMed: <https://pubmed.ncbi.nlm.nih.gov/30760447/>  
 PubMed Central: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6372921/>
9. Doust J, Glasziou P. Monitoring in clinical biochemistry. *Clin Biochem Rev.* 2013 Aug;34(2):85-92.  
 View at: PubMed: <https://pubmed.ncbi.nlm.nih.gov/24151344/>  
 PubMed Central: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3799222/>
  10. Diagnostic Errors: Technical Series on Safer Primary Care. Geneva: World Health Organization; 2016. Licence: CC BY-NC-SA 3.0 IGO  
 View at: Publisher Site: <https://www.who.int/publications/i/item/diagnostic-errors>  
 URL: <https://apps.who.int/iris/bitstream/handle/10665/252410/9789241511636-eng.pdf>
  11. Newman-Toker DE, Moy E, Valente E, Coffey R, Hines AL. Missed diagnosis of stroke in the emergency department: a cross-sectional analysis of a large population-based sample. *Diagnosis (Berl).* 2014 Jun;1(2):155-66. DOI: 10.1515/dx-2013-0038.  
 View at: Publisher Site: <https://www.degruyter.com/document/doi/10.1515/dx-2013-0038/html>  
 PubMed: <https://pubmed.ncbi.nlm.nih.gov/28344918/>  
 PubMed Central: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5361750/>
  12. Tarnutzer AA, Lee SH, Robinson KA, Wang Z, Edlow JA, Newman-Toker DE. ED misdiagnosis of cerebrovascular events in the era of modern neuroimaging: A meta-analysis. *Neurology.* 2017 Apr 11;88(15):1468-77. DOI: 10.1212/WNL.00000000000003814  
 View at: Publisher Site: <https://n.neurology.org/content/88/15/1468>  
 PubMed: <https://pubmed.ncbi.nlm.nih.gov/28356464/>  
 PubMed Central: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5386439/>
  13. Zhang Y, Xia H, Wang Y, Chen L, Li S, Hussein IA, Wu Y, Shang Y, Yao S, Du R. The rate of missed diagnosis of lower-limb DVT by ultrasound amounts to 50% or so in patients without symptoms of DVT: A meta-analysis. *Medicine (Baltimore).* 2019 Sep;98(37):e17103. DOI: 10.1097/MD.00000000000017103.  
 View at: Publisher Site: [https://journals.lww.com/md-journal/Fulltext/2019/09130/The\\_rate\\_of\\_missed\\_diagnosis\\_of\\_lower\\_limb\\_DVT\\_by.32.aspx](https://journals.lww.com/md-journal/Fulltext/2019/09130/The_rate_of_missed_diagnosis_of_lower_limb_DVT_by.32.aspx)  
 PubMed: <https://pubmed.ncbi.nlm.nih.gov/31517841/>  
 PubMed Central: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6750306/>
  14. Walen S, Damoiseaux RA, Uil SM, van den Berg JW. Diagnostic delay of pulmonary embolism in primary and secondary care: a retrospective cohort study. *Br J Gen Pract.* 2016 Jun;66(647):e444-50. DOI: 10.3399/bjgp16X685201.  
 View at: Publisher Site: <https://bjgp.org/content/66/647/e444>  
 PubMed: <https://pubmed.ncbi.nlm.nih.gov/27114207/>  
 PubMed Central: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4871310/>
  15. Michaels AD, Spinler SA, Leeper B, Ohman EM, Alexander KP, Newby LK, Ay H, Gibler WB; American Heart Association Acute Cardiac Care Committee of the Council on Clinical Cardiology, Council on Quality of Care and Outcomes Research; Council on Cardiopulmonary, Critical Care, Perioperative, and Resuscitation; Council on Cardiovascular Nursing; Stroke Council. Medication errors in acute cardiovascular and stroke patients: a scientific statement from the American Heart Association. *Circulation.* 2010 Apr 13;121(14):1664-82. DOI: 10.1161/CIR.0b013e3181d4b43e.  
 View at: Publisher Site: <https://www.ahajournals.org/doi/10.1161/CIR.0b013e3181d4b43e>  
 PubMed: <https://pubmed.ncbi.nlm.nih.gov/20308619/>
  16. Muroi M, Shen JJ, Angosta A. Association of medication errors with drug classifications, clinical units, and consequence of errors: Are they related? *Appl Nurs Res.* 2017 Feb;33:180-5. DOI: 10.1016/j.apnr.2016.12.002.  
 View at: Scopus: <https://www.sciencedirect.com/science/article/pii/S0897189716303767?via%3Dihub>  
 PubMed: <https://pubmed.ncbi.nlm.nih.gov/28096015/>
  17. Gelchu T, Abdela J. Drug therapy problems among patients with cardiovascular disease admitted to the medical ward and had a follow-up at the ambulatory clinic of Hiwot Fana Specialized University Hospital: The case of a tertiary hospital in eastern Ethiopia. *SAGE Open Med.* 2019 Jul 18;7:2050312119860401. DOI: 10.1177/2050312119860401.  
 View at: Publisher Site: <https://journals.sagepub.com/doi/10.1177/2050312119860401>  
 PubMed: <https://pubmed.ncbi.nlm.nih.gov/31367379/>  
 PubMed Central: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6643177/>
  18. Reeve A, Simcox E, Turnbull D. Ageing and Parkinson's disease: why is advancing age the biggest risk factor? *Ageing Res Rev.* 2014 Mar;14(100):19-30. DOI: 10.1016/j.arr.2014.01.004.  
 View at: Scopus: <https://www.sciencedirect.com/science/article/pii/S1568163714000051?via%3Dihub>  
 PubMed: <https://pubmed.ncbi.nlm.nih.gov/24503004/>  
 PubMed Central: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3989046/>
  19. Rees RN, Acharya AP, Schrag A, Noyce AJ. An early diagnosis is not the same as a timely diagnosis of Parkinson's disease. *F1000Res.* 2018 Jul 18;7:F1000 Faculty Rev-1106. DOI: 10.12688/f1000research.14528.1.eCollection 2018  
 View at: Publisher Site: <https://f1000research.com/articles/7-1106/v1>  
 PubMed: <https://pubmed.ncbi.nlm.nih.gov/30079229/>  
 PubMed Central: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6053699/>



20. Emamzadeh FN, Surguchov A. Parkinson's Disease: Biomarkers, Treatment, and Risk Factors. *Front Neurosci.* 2018 Aug 30;12:612. DOI: 10.3389/fnins.2018.00612.  
View at: Publisher Site: <https://www.frontiersin.org/articles/10.3389/fnins.2018.00612/full>  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/30214392/>  
PubMed Central: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6125353/>
21. Obeso JA, Stamelou M, Goetz CG, Poewe W, Lang AE, Weintraub D, Burn D, Halliday GM, Bezard E, Przedborski S, Lehericy S, Brooks DJ, Rothwell JC, Hallett M, DeLong MR, Marras C, Tanner CM, Ross GW, Langston JW, Klein C, Bonifati V, Jankovic J, Lozano AM, Deuschl G, Bergman H, Tolosa E, Rodriguez-Violante M, Fahn S, Postuma RB, Berg D, Marek K, Standaert DG, Surmeier DJ, Olanow CW, Kordower JH, Calabresi P, Schapira AHV, Stoessl AJ. Past, present, and future of Parkinson's disease: A special essay on the 200th Anniversary of the Shaking Palsy. *Mov Disord.* 2017 Sep;32(9):1264-310. DOI: 10.1002/mds.27115.  
View at: Publisher Site: <https://movementdisorders.onlinelibrary.wiley.com/doi/10.1002/mds.27115>  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/28887905/>  
PubMed Central: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5685546/>
22. Adler CH, Beach TG, Hentz JG, Shill HA, Caviness JN, Driver-Dunckley E, Sabbagh MN, Sue LI, Jacobson SA, Belden CM, Dugger BN. Low clinical diagnostic accuracy of early vs advanced Parkinson disease: clinicopathologic study. *Neurology.* 2014 Jul 29;83(5):406-12. DOI: 10.1212/WNL.0000000000000641.  
View at: Publisher Site: <https://n.neurology.org/content/83/5/406>  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/24975862/>  
PubMed Central: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4132570/>
23. Lane M, Yadav V. Multiple Sclerosis. *Textbook of Natural Medicine.* 2020:1587-1599.e3. DOI: 10.1016/B978-0-323-43044-9.00199-0.  
View at: Scopus: <https://www.sciencedirect.com/science/article/pii/B9780323430449001990?via%3Dihub>  
PubMed Central: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7348625/>
24. Culpepper WJ, Marrie RA, Langer-Gould A, Wallin MT, Campbell JD, Nelson LM, Kaye WE, Wagner L, Tremlett H, Chen LH, Leung S, Evans C, Yao S, LaRocca NG; United States Multiple Sclerosis Prevalence Workgroup (MSPWG). Validation of an algorithm for identifying MS cases in administrative health claims datasets. *Neurology.* 2019 Mar 5;92(10):e1016-e1028. DOI: 10.1212/WNL.0000000000007043.  
View at: Publisher Site: <https://n.neurology.org/content/92/10/e1016>  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/30770432/>  
PubMed Central: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6442008/>
25. Mackenzie IS, Morant SV, Bloomfield GA, MacDonald TM, O'Riordan J. Incidence and prevalence of multiple sclerosis in the UK 1990-2010: a descriptive study in the General Practice Research Database. *Journal of Neurology, Neurosurgery & Psychiatry.* 2014;85:76-84. DOI: 10.1136/jnnp-2013-305450  
View at: Publisher Site: <https://jnnp.bmj.com/content/85/1/76>  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/24052635/>  
PubMed Central: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3888639/>
26. Ziemssen T, Akgün K, Brück W. Molecular biomarkers in multiple sclerosis. *J Neuroinflammation.* 2019 Dec 23;16(1):272. DOI: 10.1186/s12974-019-1674-2.  
View at: Publisher Site: <https://jneuroinflammation.biomedcentral.com/articles/10.1186/s12974-019-1674-2>
27. Thompson AJ, Banwell BL, Barkhof F, Carroll WM, Coetzee T, Comi G, Correale J, Fazekas F, Filippi M, Freedman MS, Fujihara K, Galetta SL, Hartung HP, Kappos L, Lublin FD, Marrie RA, Miller AE, Miller DH, Montalban X, Mowry EM, Sorensen PS, Tintoré M, Traboulsee AL, Trojano M, Uitdehaag BMJ, Vukusic S, Waubant E, Weinshenker BG, Reingold SC, Cohen JA. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol.* 2018 Feb;17(2):162-73. DOI: 10.1016/S1474-4422(17)30470-2.  
View at: Publisher Site: [https://www.thelancet.com/journals/lanneur/article/PIIS1474-4422\(17\)30470-2/fulltext](https://www.thelancet.com/journals/lanneur/article/PIIS1474-4422(17)30470-2/fulltext)  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/29275977/>
28. Kaisey M, Solomon AJ, Luu M, Giesser BS, Sicotte NL. Incidence of multiple sclerosis misdiagnosis in referrals to two academic centers. *Mult Scler Relat Disord.* 2019 May;30:51-6. DOI: 10.1016/j.msard.2019.01.048.  
View at: Publisher Site: [https://www.msard-journal.com/article/S2211-0348\(19\)30048-3/fulltext](https://www.msard-journal.com/article/S2211-0348(19)30048-3/fulltext)  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/30738280/>
29. Solomon AJ. Diagnosis, Differential Diagnosis, and Misdiagnosis of Multiple Sclerosis. *Continuum (Minneapolis Minn).* 2019 Jun;25(3):611-35. DOI: 10.1212/CON.0000000000000728.  
View at: Publisher Site: [https://journals.lww.com/continuum/Abstract/2019/06000/Diagnosis,\\_Differential\\_Diagnosis,\\_and.5.aspx](https://journals.lww.com/continuum/Abstract/2019/06000/Diagnosis,_Differential_Diagnosis,_and.5.aspx)  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/31162308/>
30. 2020 Alzheimer's disease facts and figures. *Alzheimers Dement.* 2020 Mar 10. DOI: 10.1002/alz.12068.  
View at: Publisher Site: <https://alz-journals.onlinelibrary.wiley.com/doi/10.1002/alz.12068>  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/32157811/>
31. World Health Organization. Global action plan on the public health response to dementia 2017-2025. Geneva

2017. Licence: CC BY-NC-SA 3.0 IGO.  
View at: Publisher Site: <https://apps.who.int/iris/handle/10665/259615>  
URL: <https://apps.who.int/iris/bitstream/handle/10665/259615/9789241513487-eng.pdf?sequence=1>
32. Bloudek LM, Spackman DE, Blankenburg M, Sullivan SD. Review and meta-analysis of biomarkers and diagnostic imaging in Alzheimer's disease. *J Alzheimers Dis*. 2011;26(4):627-45. DOI: 10.3233/JAD-2011-110458.  
View at: Publisher Site: <https://content.iospress.com/articles/journal-of-alzheimers-disease/jad110458>  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/21694448/>
33. McKhann GM, Albert MS, Sperling RA. Changing diagnostic concepts of Alzheimer's disease. In: Hampel H, Carrillo MC, eds. *Alzheimer's disease – Modernizing concept, biological diagnosis and therapy*. Basel, Switzerland: Karger; 2012: p. 115–21.  
View at: Publisher Site: <https://www.karger.com/Book/Home/256724>  
Scielo: [https://scielo.isciii.es/scielo.php?script=sci\\_arttext&pid=S0213-61632013000200007](https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S0213-61632013000200007)
34. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin*. 2020 Jan;70(1):7-30. DOI: 10.3322/caac.21590.  
View at: Publisher Site: <https://acsjournals.onlinelibrary.wiley.com/doi/10.3322/caac.21590>
35. Morris LG, Tuttle RM, Davies L. Changing Trends in the Incidence of Thyroid Cancer in the United States. *JAMA Otolaryngol Head Neck Surg*. 2016 Jul 1;142(7):709-11. DOI: 10.1001/jamaoto.2016.0230.  
View at: Publisher Site: <https://jamanetwork.com/journals/jamaotolaryngology/fullarticle/2513194>  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/27078686/>  
PubMed Central: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4956490/>
36. Ahn HS, Kim HJ, Welch HG. Korea's thyroid-cancer "Epidemic" – screening and overdiagnosis. *N Engl J Med*. 2014 Nov 6;371(19):1765-7. DOI: 10.1056/NEJMp1409841.  
View at: Publisher Site: <https://www.nejm.org/doi/10.1056/NEJMp1409841>  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/25372084/>
37. Ahn HS, Welch HG. South Korea's Thyroid-Cancer "Epidemic" – Turning the Tide. *N Engl J Med*. 2015 Dec 10;373(24):2389-90. DOI: 10.1056/NEJMc1507622.  
View at: Publisher Site: <https://www.nejm.org/doi/10.1056/NEJMc1507622>  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/26650173/>
38. Theoharis CG, Schofield KM, Hammers L, Udelsman R, Chhieng DC. The Bethesda thyroid fine-needle aspiration classification system: year 1 at an academic institution. *Thyroid*. 2009 Nov;19(11):1215-23. DOI: 10.1089/thy.2009.0155.  
View at: Publisher Site: <https://www.liebertpub.com/doi/10.1089/thy.2009.0155>  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/19888859/>
39. Wang W, Chang J, Jia B, Liu J. The Blood Biomarkers of Thyroid Cancer. *Cancer Manag Res*. 2020 Jul 6;12:5431-8. DOI: 10.2147/CMAR.S261170.  
View at: Publisher Site: <https://www.dovepress.com/the-blood-biomarkers-of-thyroid-cancer-peer-reviewed-fulltext-article-CMAR>
40. Ilic D, Neuberger MM, Djulbegovic M, Dahm P. Screening for prostate cancer. *Cochrane Database Syst Rev*. 2013 Jan 31;(1):CD004720. DOI: 10.1002/14651858.CD004720.pub3.  
View at: Publisher Site: <https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD004720.pub3/abstract>  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/23440794/>
41. Prabhu V, Lee T, McClintock TR, Lepor H. Short-, Intermediate-, and Long-term Quality of Life Outcomes Following Radical Prostatectomy for Clinically Localized Prostate Cancer. *Rev Urol*. 2013;15(4):161-77.  
View at: PubMed: <https://pubmed.ncbi.nlm.nih.gov/24659913/>  
PubMed Central: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3922321/>
42. Qu M, Ren SC, Sun YH. Current early diagnostic biomarkers of prostate cancer. *Asian J Androl*. 2014 Jul-Aug;16(4):549-54. DOI: 10.4103/1008-682X.129211.  
View at: Publisher Site: <https://www.ajandrology.com/article.asp?issn=1008-682X;year=2014;volume=16;issue=4;spage=549;epage=554;aulast=Qu>  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/24830695/>  
PubMed Central: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4104079/>
43. Stenvinkel P. Chronic kidney disease: a public health priority and harbinger of premature cardiovascular disease. *J Intern Med*. 2010 Nov;268(5):456-67. DOI: 10.1111/j.1365-2796.2010.02269.x.  
View at: Publisher Site: <https://onlinelibrary.wiley.com/doi/10.1111/j.1365-2796.2010.02269.x>  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/20809922/>
44. Uchino S, Kellum JA, Bellomo R, Doig GS, Morimatsu H, Morgera S, Schetz M, Tan I, Bouman C, Macedo E, Gibney N, Tolwani A, Ronco C; Beginning and Ending Supportive Therapy for the Kidney (BEST Kidney) Investigators. Acute renal failure in critically ill patients: a multinational, multicenter study. *JAMA*. 2005 Aug 17;294(7):813-8. DOI: 10.1001/jama.294.7.813.  
View at: Publisher Site: <https://jamanetwork.com/journals/jama/fullarticle/201386>  
PubMed: <https://pubmed.ncbi.nlm.nih.gov/16106006/>
45. Bosch JP. Renal reserve: a functional view of glomerular filtration rate. *Semin Nephrol*. 1995 Sep;15(5):381-5.  
View at: PubMed: <https://pubmed.ncbi.nlm.nih.gov/8525139/>
46. Herrera J, Rodríguez-Iturbe B. Stimulation of tubular secretion of creatinine in health and in conditions as-

- sociated with reduced nephron mass. Evidence for a tubular functional reserve. *Nephrol Dial Transplant*. 1998 Mar;13(3):623-9. DOI: 10.1093/ndt/13.3.623. View at: Academic: <https://academic.oup.com/ndt/article/13/3/623/1848120>
47. Glasscock RJ, Denic A, Rule AD. The conundrums of chronic kidney disease and aging. *J Nephrol*. 2017 Aug;30(4):477-83. DOI: 10.1007/s40620-016-0362-x. View at: Scopus: <https://link.springer.com/article/10.1007%2Fs40620-016-0362-x> PubMed: <https://pubmed.ncbi.nlm.nih.gov/27885585/>
48. Jiang J, Li X, Zhao C, Guan Y, Yu Q. Learning and inference in knowledge-based probabilistic model for medical diagnosis. *Knowledge-Based Systems*. 2017;138: 58-68, ISSN 0950-7051. DOI: 10.1016/J.KNOSYS.2017.09.030. View at: Scopus: <https://www.sciencedirect.com/science/article/abs/pii/S0950705117304495?via%3Dihub> Publisher Site: <https://dl.acm.org/doi/abs/10.1016/j.knosys.2017.09.030> Booksc: <https://ur.booksc.eu/book/67269701/cc2c8c>
49. Seh AH, Zarour M, Alenezi M, Sarkar AK, Agrawal A, Kumar R, Khan RA. *Healthcare Data Breaches: Insights and Implications*. Healthcare (Basel). 2020 May 13;8(2):133. DOI: 10.3390/healthcare8020133. View at: Publisher Site: <https://www.mdpi.com/2227-9032/8/2/133> PubMed: <https://pubmed.ncbi.nlm.nih.gov/32414183/> PubMed Central: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7349636/>
50. Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. 1995; 20: 273-97. DOI: 10.1007/BF00994018 View at: Scopus: <https://link.springer.com/article/10.1007/BF00994018> URL: <https://link.springer.com/content/pdf/10.1007/bf00994018.pdf>
51. McCallum A, Nigam K. A comparison of event models for naive bayes text classification. In: AAAI-98 workshop on learning for text categorization. 1998, July;752(1): 41-8. View at: Bibsonomy: <https://www.bibsonomy.org/bibtex/2fa46d1cc0dd56ab40a7f722e569a1fd3/jil> URL: <https://www.cs.cmu.edu/~knigam/papers/multi-nomial-aaaiws98.pdf>

Article history  
Received: 18.04.2021  
Revision requested: 02.06.2021  
Revision received: 29.07.2021  
Accepted: 24.09.2021  
Published: 30.09.2021

## ВПРОВАДЖЕННЯ АЛГОРИТМІВ МАШИННОГО НАВЧАННЯ В СИСТЕМІ ОХОРОНИ ЗДОРОВ'Я ЯК ПЕРСПЕКТИВНИЙ НАПРЯМОК ДЛЯ НАУКИ, ОХОРОНИ ЗДОРОВ'Я ТА БІЗНЕСУ

*Василевський В., Степанов І., Коваль Р., Сопутняк М., Лютянська Н., Шейко В., Ставничий Т.*

*Товариство з обмеженою відповідальністю BIONITY, Київ, Україна*

*byblikpeper@gmail.com*

**Актуальність.** Сучасний стан медицини є недосконалим. Серед основних проблем можна виділити діагностику і лікування захворювань. Труднощі роботи з біомедичними даними є серйозним обмежуючим фактором у вирішенні найважливіших завдань охорони здоров'я, які представлені в статистично достовірних групах захворювань. Велика кількість даних в природничих науках створює як можливості, так і проблеми для їх ефективного використання в клінічній практиці. Інструменти на основі машинного навчання необхідні для генерації нових ідей і виявлення нових прихованих закономірностей, особливо при великих наборах даних. Рішення на основі штучного інтелекту можуть успішно використовуватися для діагностики захворювань, моніторингу загального стану здоров'я, прогнозування ризиків, прийняття рішень для лікування і отримання біомедичних знань.

**Ціль:** проаналізувати потенціал алгоритмів машинного навчання в охороні здоров'я на існуючих проблемах і зробити прогноз їх розвитку в найближчому майбутньому.

**Метод.** Аналітичний огляд літератури по ключовим словам із наукометричних баз даних Scopus, PubMed, Wiley. Глибина пошуку 7 років з 2013 по 2020 рік.

**Результати.** Аналізуючи сучасний загальний стан системи охорони здоров'я, ми виділили найбільш актуальні проблеми, пов'язані з діагностикою, лікуванням і системним управлінням: помилки діагностики, пізня діагностика (в тому числі під час надзвичайних ситуацій), гіпердіагностика, бюрократія, проблеми з комунікацією і труднощі передачі даних. Ми вивчили деталі процесу, зручного для прийняття рішень в клінічному середовищі, для того, щоб визначити точні моменти, які можуть бути значно поліпшені за допомогою рішень, прийнятих на основі штучного інтелекту, в тому числі: діагностика захворювань, моніторинг загального стану здоров'я, прогнозування ризиків, рішення про лікування і отримання біомедичних знань. Ми позначили алгоритми машинного навчання як перспективний інструмент для діагностики та лікування захворювань, а також для генерації нових корисних ідей і обробки великих наборів даних.

**Висновок.** Машинне навчання – це група технологій, які можуть стати основним фактором для вирішення різних медичних проблем. Але все ж є ряд проблем, які потрібно вирішити перед інтенсивним впровадженням таких інструментів в систему охорони здоров'я.

**Ключові слова:** машинне навчання, біомедичні дані, проблеми охорони здоров'я, обробка біомедичних даних

**ВНЕДРЕНИЕ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ В СИСТЕМЕ ЗДРАВООХРАНЕНИЯ КАК ПЕРСПЕКТИВНОЕ НАПРАВЛЕНИЕ ДЛЯ НАУКИ, ЗДРАВООХРАНЕНИЯ И БИЗНЕСА**

*Василевский В., Степанов И., Коваль Р., Сопутняк М., Лютянская Н., Шейко В., Ставничий Т.*

*Общество с ограниченной ответственностью BIONITY, Киев, Украина*

byblikpeper@gmail.com

**Актуальность.** Современное состояние медицины несовершенно. Среди основных проблем можно выделить диагностику и лечение заболеваний. Трудности работы с биомедицинскими данными являются серьезным ограничивающим фактором в решении важнейших задач здравоохранения, которые представлены в статистически достоверных группах заболеваний. Большое количество данных в естественных науках создает как возможности, так и проблемы для их эффективного использования в клинической практике. Инструменты на основе машинного обучения необходимы для генерации новых идей и обнаружения новых скрытых закономерностей, особенно при больших наборах данных. Решения на основе искусственного интеллекта могут успешно использоваться для диагностики заболеваний, мониторинга общего состояния здоровья, прогнозирования рисков, принятия решений для лечения и получения биомедицинских знаний.

**Цель:** проанализировать потенциал алгоритмов машинного обучения в здравоохранении на существующих проблемах и сделать прогноз их развития в ближайшем будущем.

**Метод.** Аналитический обзор литературы по ключевым словам из наукометрических баз данных Scopus, PubMed, Wiley. Глубина поиска 7 лет с 2013 по 2020.

**Результаты.** Анализируя современное общее состояние системы здравоохранения, мы выделили наиболее актуальные проблемы, связанные с диагностикой, лечением и системным управлением: ошибки диагностики, поздняя диагностика (в том числе во время чрезвычайных ситуаций), гипердиагностика, бюрократия, проблемы с коммуникацией и трудности передачи данных. Мы изучили детали процесса, удобного для принятия решений в клинической бреде, для того, чтобы определить точные моменты, которые могут быть значительно улучшены с помощью решений, принятых на основе искусственного интеллекта, в том числе: диагностика заболеваний, мониторинг общего состояния здоровья, прогнозирование рисков, решения о лечении, и получение биомедицинских знаний. Мы обозначили алгоритмы машинного обучения как перспективный инструмент для диагностики и лечения заболеваний, а также для генерации новых полезных идей и обработки больших наборов данных.

**Вывод.** Машинное обучение – это группа технологий, которые могут стать основным фактором для решения различных медицинских проблем. Но все же есть ряд проблем, которые нужно решить перед интенсивным внедрением таких инструментов в систему здравоохранения.

**Ключевые слова:** машинное обучение, биомедицинские данные, проблемы здравоохранения, обработка биомедицинских данных