IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, July 2023

AI Based Disease Prediction

Tulshi Sanjay Phadatare and Pooja Naik
PG Students, Masters in Computer Application
Late Bhausaheb Hiray S.S. Trust's Institute of Computer Application, Mumbai, India

Abstract: *Healthcare has benefited greatly from artificial intelligence (AI), which has revolutionized disease prediction and enhanced patient outcomes. The significance, methodology, and difficulties of AI-based disease prediction are highlighted in this paper's thorough discussion. The significance of disease prediction in healthcare is examined in the first section. Improved patient care, lower healthcare costs, and more effective public health actions all depend on early diagnosis and precise illness prognostication. With the help of massive amounts of data, AI approaches like machine learning and deep learning have demonstrated tremendous potential for identifying intricate patterns and predicting diseases. The significance of disease prediction in healthcare is examined in the first section. The approaches utilized for AI-based disease prediction are covered in detail in the second part. The difficulties in using AI to predict diseases are highlighted in the third section. The possible advantages of AI-based disease prediction are covered in the fourth part. Healthcare systems may boost preventive measures, create individualized treatment plans, distribute resources more effectively, and enhance patient outcomes by utilizing AI. Additionally, AI-based disease prediction can support public health policy, identify high-risk populations, and enable early interventions. In conclusion, AI-based disease prediction has become a viable strategy to revolutionize healthcare. Making informed judgements with the use of cutting-edge AI approaches can assist healthcare professionals and policymakers, ultimately improving patient care, illness management, and population health.*

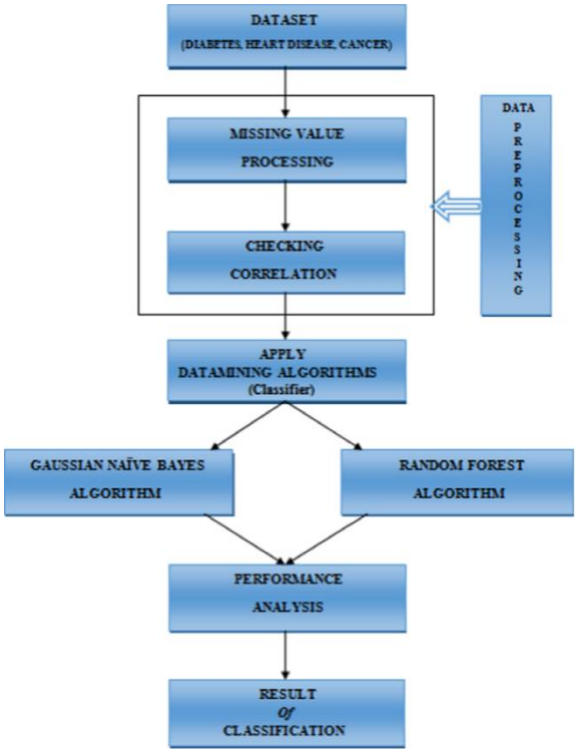
Keywords: Healthcare

I. INTRODUCTION

Artificial intelligence (AI)-based disease prediction is the application of algorithms and methodologies to forecast the onset, course, or advancement of different diseases in people. In order to find patterns and trends in huge datasets, including medical records, genetic data, lifestyle data, and other pertinent health indicators, it makes use of machine learning and data mining techniques.AI-based illness prediction aims to help medical professionals make diagnosis, prognoses, and treatment plans that are more accurate, thereby improving patient outcomes. AI systems can find subtle correlations and signs by analyzing enormous amounts of data that may not be immediately obvious to human observers. This makes it possible for early illness identification, individualized risk assessments, and proactive therapies, potentially lowering rates of morbidity and mortality. A wide variety of diseases, including but not limited to neurological ailments, cancer, respiratory disorders, cardiovascular diseases, and infectious diseases, can be predicted by AI models. These models incorporate information from a variety of sources, including genetic data, clinical examination results, laboratory test results, patient demographics, and medical history. AI algorithms may produce forecasts and risk ratings by taking into account several variables and their intricate relationships, which can help with illness prevention, early intervention, and treatment planning.

The advantages of AI-based disease prediction go beyond the treatment of specific patients. By analyzing population-level data and locating illness patterns, outbreaks, and potential risk factors, it can also support public health activities. This data can help healthcare organizations and policymakers allocate resources, create targeted solutions, and put them into practice. It's crucial to keep in mind, too, that AI-based disease prediction is intended to supplement rather than replace the knowledge of healthcare experts. When clinicians are making diagnostic and treatment decisions, the predictions made by AI models act as decision support tools by giving them more information and insights. To preserve patient privacy, fairness, and trust in the technology, meticulous validation, openness, and ethical concerns are necessary for an effective use of AI in healthcare.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, July 2023



In summary, AI-based disease prediction has the potential to completely transform healthcare by utilizing the capabilities of artificial intelligence to analyze massive volumes of data and help healthcare practitioners make more precise predictions and judgement. It has the potential to greatly enhance patient outcomes and contribute to public health by making early detection and personalized therapies possible.

II. LITERATURE REVIEW

I can provide you with a general literature analysis on AI-based disease prediction up until that point even though I don't currently have direct access to the most recent research articles because my training only lasts until September 2021. From the extant literature, the following are some salient points and trends:

Cardiovascular Diseases: By examining risk factors, medical imaging, and patient data, AI has shown potential in the prediction of cardiovascular diseases. Studies have shown how AI systems can be used to foresee outcomes like heart attacks, strokes, and cardiac events.

- Cancer: Numerous cancer forms, including lung, breast, prostate, and skin cancer, have been thoroughly investigated in terms of AI-based disease prediction. In order to predict cancer risk, tumor progression, and treatment outcomes, machine learning algorithms have been created to analyze clinical data, genetic profiles, radiological images, and histological data.
- Diabetes: AI algorithms have been used to forecast the development and course of the disease as well as its complications. These models give tailored risk assessments and intervention methods by using data from the patient's demographics, genetics, lifestyle, and glucose monitoring.
- Disorders of the nervous system: AI-based illness prediction has been used in neurology to forecast the onset and course of diseases like Alzheimer's, Parkinson's, and multiple sclerosis. To offer early detection and prognostic insights, machine learning algorithms analyze neuroimaging data, genetic information, cognitive tests, and clinical data.AI has been used in the prediction and monitoring of infectious diseases, including epidemics and outbreaks. AI models can be used to predict potential hotspots, track the transmission of diseases, and help with resource allocation by analyzing a variety of data sources, including social media, electronic health records, and syndromic monitoring systems.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, July 2023

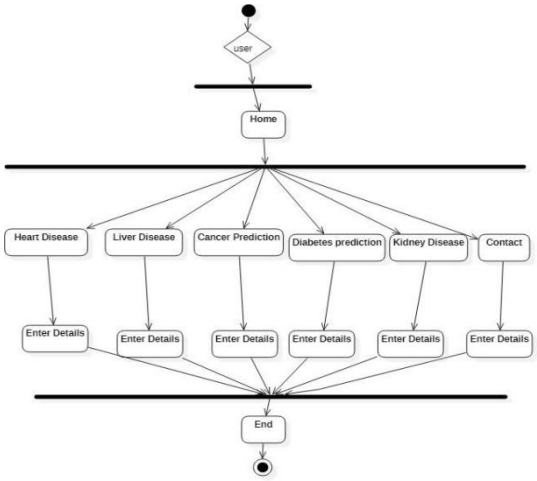
- Integration of Multiple Data Sources: To enhance illness prediction models, numerous studies have concentrated on integrating various data sources, such as electronic health records, genetics, wearable technology, and medical imaging. The risk and course of diseases can be better understood by combining several forms of data.
- Deep Learning Approaches: Due to its capacity to automatically extract pertinent characteristics from big datasets, deep learning approaches, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have become increasingly prominent in the field of disease prediction These models have produced encouraging outcomes in a number of medical fields.
- Ethical and Regulatory Considerations: When deploying AI-based disease prediction models, researchers have emphasized the significance of taking ethical issues such patient privacy, data security, openness, and fairness into account.

You can find the most recent details and developments in AI-based disease prediction by consulting the most recent scholarly publications and journals in the area.

III. PROBLEM DEFINITION

The definition of the issue for AI-based disease prediction is the use of artificial intelligence methods to precisely forecast the onset, course, or result of various diseases. The objective is to use AI algorithms and models to analyze massive datasets and derive insightful information that might help medical practitioners make better decisions about illness diagnosis, treatment, and prevention.

The following factors can help to further define the issue:



- Data Acquisition and Integration: Getting thorough and pertinent data for disease prediction is the first obstacle. In order to do this, data must be gathered from a variety of sources, including electronic health records, medical imaging, genetic data, lifestyle data, and patient-reported data. After the data has been collected, the following stage is to extract and choose significant features that can accurately describe the traits of the condition.
- Training and Validation: The chosen AI models must be trained on the data at hand to discover the underlying correlations and patterns. To attain the greatest performance, this entails dividing the dataset into training, validation, and testing sets and optimizing the model parameters. To achieve generalizability, the models need to be systematically assessed using the right performance criteria and validated on external datasets.
- Interpretability and Explain ability: Understanding the outcomes and offering justifications for the forecasts is one of the difficulties with AI-based disease prediction. To develop confidence in the AI models and make wise decisions, healthcare workers must comprehend the reasoning behind the forecasts.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, July 2023

- Deployment and Integration: Introducing AI-based illness prediction models into clinical practice is the ultimate challenge. To deliver real-time forecasts and aid healthcare professionals in their decision-making processes, the models must be incorporated into the workflows and systems already in place in the industry.

By overcoming these obstacles and creating precise and trustworthy AI-based disease prediction models, we can improve patient care by promoting early identification, individualized treatment strategies, and better health outcomes. To win the trust of both healthcare professionals and patients, it is crucial to assure the transparency, interpretability, and ethical use of AI models in healthcare.

IV. OBJECTIVE AND SCOPE

Utilizing artificial intelligence techniques and algorithms to increase the precision and effectiveness of disease prediction in healthcare is the goal of AI-based disease prediction.

The main objective is to help medical professionals make better judgments about disease diagnosis, prognosis, treatment planning, and preventive actions by providing them with more information.

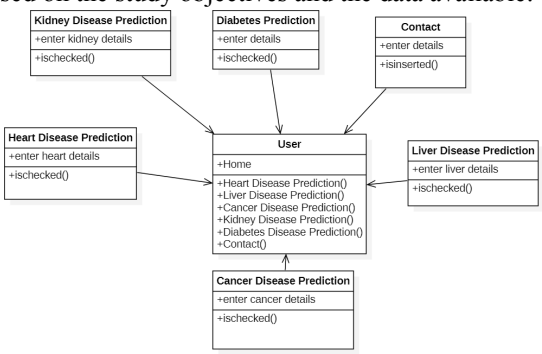
- Early Detection: To find early symptoms and risk factors for diseases, AI models can examine a variety of data sources, including patient demographics, medical history, genetic data, lifestyle factors, and medical imaging.
- Risk assessment: Using a person's features, genetic predispositions, and lifestyle factors, AI algorithms can determine their risk of contracting a certain disease.
- Prognosis and Treatment Planning: AI-based illness prediction can offer information on the likelihood that a disease will progress and how it will react to various treatments. Healthcare providers can use this data to customize treatment strategies and actions for each patient.
- Prognosis and Treatment Planning: AI-based illness prediction can offer information on the likelihood that a disease will progress and how it will react to various treatments.
- Public health and population-level analysis: AI can examine huge datasets to spot trends in disease, epidemics, and potential risk factors at the population level.
- Integration of Multiple Data Sources: AI-based disease prediction is capable of integrating a variety of data sources, such as wearable technology, genetic data, electronic health records, and medical imaging. A more thorough picture of illness risk, progression, and treatment response can be attained by merging these various data sets.

AI models can be used as decision support tools for healthcare practitioners, giving them new information and insights to improve the accuracy of their diagnoses and treatment selections. This can help with patient management, diagnostic error reduction, and providing more individualized and accurate healthcare.

Overall, improving disease prediction accuracy, enabling early detection, improving treatment planning, and supporting public health initiatives are the goals and scope of AI-based disease prediction. This will ultimately result in better patient outcomes and more effective healthcare delivery.

V. RESEARCH METHODOLOGY

The development and evaluation of predictive models using artificial intelligence methods are part of the research methodology for AI-based disease prediction. Here are the main steps that are commonly taken in the research process, but specific approaches can vary based on the study objectives and the data available:



IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, July 2023

- **Data Gathering:** The initial step in the disease prediction task is to gather pertinent data. Obtaining genetic profiles, lifestyle data, medical imaging data, electronic health records, and other pertinent sources may be necessary for this. Ethics and privacy laws should be followed during the data collection procedure.
- **Data preprocessing** is necessary to ensure the consistency and quality of the data after it has been collected. As part of this, the data must be cleaned, handled for missing values, outliers eliminated, and the data must be normalized or standardized. The most informative features are selected from the available dataset in this step's feature selection process. The most important features that support the disease prediction task can be determined using feature selection approaches like statistical testing, correlation analysis, or dimensionality reduction techniques.
- **Model Selection**: The next step entails picking the best AI models and algorithms for predicting diseases. This decision is based on the particular research question and the type of information that is at hand. Convolutional neural networks (CNNs) or recurrent neural networks (RNNs) are examples of common AI models used in disease prediction.
- **Model Training and Evaluation:** Using the given dataset, the chosen AI models are trained. Typically, training, validation, and testing sets are separated from the dataset. model's settings and discover the underlying
- **Evaluation of Performance:** The prediction models are assessed using a variety of performance indicators. Accuracy, precision, recall, F1-score, area under the receiver operating characteristic curve (AUC-ROC), and other domain-specific metrics are a few examples of these metrics.
- **Interpretability and Explain ability:** The interpretability and explain ability of the AI models may be crucial, depending on the study objectives. In order to provide insight into the characteristics and elements influencing the disease prediction. This stage makes it easier for patients and healthcare professionals to comprehend the model's judgements and builds faith in its judgements.
- **Validation and External Testing:** The created models should be evaluated on external datasets or in actual clinical settings to ensure that they are robust and generalizable. External validation enables the model to be successfully deployed in real-world healthcare scenarios and to function well on fresh, untested data.
- **Ethics:** Throughout the research process, ethical issues such data security, privacy, and fairness should be carefully considered. To ensure ethical and responsible use of AI in disease prediction, regulatory criteria and guidelines must be followed.

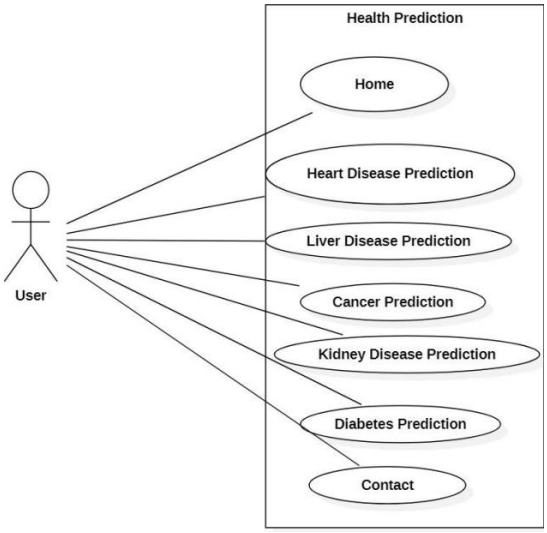
By using this study methodology, scientists may create and assess AI-based illness prediction models, enabling precise forecasts that can support initiatives for better patient care and public health.

VI. ANALYSIS AND FINDING

Depending on the exact research goals, disease area, and dataset utilized, the analysis and results of AI-based disease prediction studies can differ. However, the following analysis and conclusions are frequently reported by researchers:
Model Performance: Using a variety of evaluation criteria, including accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC), researchers evaluate the performance of AI-based illness prediction models. The conclusions shed light on the models' accuracy in forecasting disease outcomes when compared to standard or existing approaches.
The importance of various traits or factors in disease prediction models is examined by researchers.
The components that significantly contribute to the prediction task can be found out thanks to this investigation. It offers perceptions into the primary indications or risk.

- Prediction Accuracy: A key result is how well AI models anticipate the onset, course, or outcome of diseases or medical treatments. In order to provide an indication of the model's predictive power and possible clinical utility, researchers provide the success rates of the models in properly predicting positive and negative cases.
- Comparative Analysis: To find the optimum method for a given task of disease prediction, researchers frequently evaluate the performance of various AI algorithms or models.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, July 2023



- Clinical Impact: Scientists evaluate the possible clinical effects of disease prediction models based on AI. This includes assessing how the application of AI models might better patient outcomes, enable early detection, and increase diagnostic accuracy.
- Generalizability: By putting newly created models to the test on external datasets or in actual clinical settings, researchers look into the generalizability of the models.
- Interpretability and Explain ability: If these two factors were crucial study goals, researchers examine the model's interpretability procedures and outcomes. They report on how well the models can account for their results and offer details on the variables affecting the disease prognosis. The model's decisions are made more credible and easier to embrace in clinical practice thanks to this study.
- Future Directions: Researchers talk about how their AI-based disease prediction models are constrained by factors like data availability, sample size, bias, or generalizability. They also point out areas for further study and development, offering ways to improve the models.

Together, these analyses and conclusions help further our understanding of how AI might be used to forecast diseases, confirm the efficacy of the models that have been created, and direct future studies and applications in the healthcare industry.

6.1 Proposed algorithm

A. Naive Bayes classification algorithm:

The Bayes theorem was employed in this instance to classify data and make the assumption that classifying data is predictor independent. It assumes that when a certain feature is present in a class, the Naive Bayes classifier is independent of all other features.

The construction and analysis of very large datasets can both be done using the Naive Bayes model. This model is an extremely smart and simple classification system, and it excelled even in challenging situations. Calculate the posterior probability using the Bayes theory and the following equation:

$P(a/y)=(P(y/a)P(a))/P(y)$

where P(a/y) denotes the class's posterior probability, P(a) its prior probability, P(y/a) its likelihood, or the likelihood of the predictor given the class, and P(y) its prior probability.

B. Random Forest algorithm

The random forest (RF) is a hierarchical grouping of base classifiers with a tree topology. Textual information typically has several different dimensions. There are many irrelevant attributes in the dataset. For the classifier model, just a few significant features are informative. To choose the most crucial significant property, the RF algorithm uses a straightforward predefined probability. Bierman developed the RF technique by translating a random sample of feature subspaces to sample data subsets and various decision trees.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, July 2023

Random forest pseudocode:

1. Choose "n" features at random from a total of "k" features. where n << k.
2. Using the optimal split point, determine the node "n" among the "n" features.
3. Using the best split, group the node into daughter nodes.
4. Continue in steps 1 through 3 until "l" nodes have been reached.
5. Create a forest by repeating steps 1 through 4 "n" times to produce "n" trees.

VII. LIMITATIONS AND FUTURE SCOPE

Although AI-based disease prediction has advanced significantly in recent years, it still has several drawbacks. For future improvements and determining the potential use of AI-based disease prediction, it is essential to comprehend these limits. Here are some typical restrictions and probable directions for the future:

AI models rely significantly on representative, high-quality datasets. Data availability. However, it can be difficult to get extensive and varied datasets for disease prediction.

- Interpretability and Explain ability: Many AI models, including deep learning architectures, are sometimes regarded as "black boxes," making it challenging to comprehend the underlying mechanisms affecting disease outcomes and to interpret their predictions. AI models may find it difficult to generalize successfully to varied patient groups or healthcare environments if they were trained on particular datasets.
- Legal and Ethical Issues: The use of AI-based disease prediction systems poses legal and ethical issues, such as patient privacy, data security, permission, and potential algorithmic biases.
- Integration into Clinical Practice: Despite encouraging results, it is still difficult to incorporate AI-based disease prediction models into everyday clinical practice. Future research should concentrate on creating practical interfaces that are easy for clinicians to use and understand to comprehend AI predictions.
- Long-term Outcomes and Validation: AI-based illness prediction models frequently concentrate on short-term outcomes, including the onset or advancement of the disease within a certain time frame. Future studies should look into how AI predictions affect patient outcomes, treatment choices, and healthcare expenditures over the long term.
- Real-time Predictions and Continuous Monitoring: AI-based disease prediction can benefit from real-time predictions and continuous monitoring of patient data.

Overall, addressing the issues and difficulties, improving interpretability and explain ability, enhancing generalizability and ethical considerations, and smoothly incorporating AI models into clinical practice are key to the future of AI-based disease prediction.

VIII. CONCLUSION

The potential for AI-based disease prediction to transform healthcare by enhancing disease diagnosis, prognosis, treatment planning, and preventive measures has been highlighted. In order to analyze huge and complicated datasets and uncover insightful patterns that might help healthcare practitioners make better decisions, artificial intelligence algorithms and models are used.AI-based disease prediction models can give early disease detection, evaluate individual risk factors, and provide individualized treatment regimens by utilizing machine learning, deep learning, and other AI approaches.

These models may improve patient outcomes, allocate healthcare resources more effectively, and support public health programmers fully utilize the potential of AI-based disease prediction, there are a few obstacles and constraints that must be overcome.

These include concerns with data accessibility and quality, interpretability and explicability of AI models, generalizability across various demographics and healthcare contexts, ethical and regulatory challenges, integration into clinical practice, and long-term validation of results. Despite these obstacles, current investigations and improvements in AI-based disease prediction present encouraging prospects for future progress. Enhancing model interpretability and explain ability, resolving ethical and regulatory issues, and smoothly incorporating AI into hospital procedures are all possible areas for future research. To advance research, assure appropriate AI use, and realize the full promise of AI-based diseases, collaboration between AI researchers, physicians, and policymakers will be essential.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, July 2023

REFERENCES

[1] Ranjit JA, Shanmugaratnam KL (2010) Distributed and cooperative multi-agent based intrusion detection system. Indian J SciTechnol 3(10):1070–1074
[2] Wu H, Yang S, Huang Z, He J, Wang X (2018) Type 2 diabetes mellitus prediction model based on data mining. Inform Med Unlocked 10:100–107
[3] Yi Y, Wu J, Xu W (2011) Incremental SVM based on reserved set for network intrusion detection. Elsevier J Expert SystAppl 38(6):7698–7707
[4] Kumari M, Vohra R, Arora A (2014) Prediction of diabetes using Bayesian network. Int J Compute SciInfTechnol (IJCSIT) 5(4):5174–5178
[5] Long NC, Messed P, Unger H (2015) A highly accurate firefly-based algorithm for heart disease prediction. Expert SystAppl 42:8221–8231
[6] Abiyev, R. H., and Ma'aitah, M. K. S. (2018). Deep convolutional neural networks for chest diseases detection. Journal of Healthcare Engineering, 2018.
[7] Dande, P., and Samant, P. (2018). Acquaintance to artificial neural networks and use of artificial intelligence as a diagnostic tool for tuberculosis: a review. Tuberculosis, 108, 1–9.
[8] Nilashi, M., Ahmadi, H., Sheikhtaheri, A., et al. (2020). Remote tracking of Parkinson's Disease progression using ensembles of Deep Belief Network and Self-Organizing Map. Expert Systems with Applications, 159, 113562
[9] Davenport, T., and Kalakota, R. (2019). The potential for artificial intelligence in healthcare. Future Healthcare Journal, 6(2), 94.
[10] Ismaeel, S., Miri, A., and Chourishi, D. (2015, May). Using the Extreme Learning Machine (ELM) technique for heart disease diagnosis. In 2015 IEEE Canada International Humanitarian Technology Conference (IHTC2015) (pp. 1–3). IEEE.