Journal of Autonomous Intelligence (2024) Volume 7 Issue 3
doi: 10.32629/jai.v7i3.999

ORIGINAL RESEARCH ARTICLE

Study on prediction and diagnosis AI model of frequent chronic diseases based on health checkup big data

Jae Young Park[1,*], Jai-Woo Oh[2]

[1] *Department of Health Management & Education, Kyungdong University, Gangwon 26495, Korea*

[2] *Department of Health Management & Education, Kyungdong University, 15810, Korea*

*** Corresponding author:** Jae Young Park, pjy@kduniv.ac.kr

ABSTRACT

The purpose of this study is to develop a disease prediction model that can evaluate diagnostic test results based on a machine learning model and big data analysis algorithms for automated judgment of health chuck-up results. The research method used the catboost algorithm for data pretreatment and analysis. The original data was divided into learning data and test data to ensure 21,140 effective data consisting of 27 properties and to develop and utilize predictive models. Learning data was used as input data for the development of predictive models, and the test data was divided into data for the performance evaluation of the predictive model. Random forest analysis algorithms were used to analyze testing and determination accuracy that affect disease determination, and forecasting model performance analysis was analyzed by accuracy, ROC (ROC) Area, Confusion Matrix, Precision, and Recall indicators. As a result of random forest analysis, both diabetes and two -ventilation diseases were analyzed to be used as a commercial platform model by analyzing more than 90% forecast accuracy. The results of this study found that using big data analysis and machine learning, it is possible to determine and predict specific diseases based on health check-up data.

Keywords: health check-up; big data; AI (Artificial Intelligence); deep learning; disease prediction; chronic disease

ARTICLE INFO

Received: 18 July 2023
Accepted: 28 July 2023
Available online: 8 January 2024

1. Introduction

A cutting-edge technology integrated with the latest information and communication technology (ICT) are being applied in modern medicine to not only increase life expectancy but also help people stay healthy in the later years. The advances in biotechnology, including genome analysis, as well as ICT are breaking down the barrier between advanced digital technology and medicine and driving innovation in healthcare. As a result, there is now a paradigm shift in the healthcare sector from treatment-focused care to preventive, predictive, personalized, and participatory care. Modern society is characterized by increased life expectancy and population aging resulting from medical advancements and a higher standard of living. With the population aging phenomenon becoming more pronounced, the vast majority of the diseases observed today are chronic diseases, rather than infectious diseases as was the case in the past. Chronic illnesses, which arise due to a combination of multiple factors, require ongoing treatment and care, causing a heavy financial burden. Thus, in many countries, health screening services are being expanded as a means to promote public health and for prevention and early detection of chronic diseases[1]. Some medical institutions are even offering personalized health screening services to meet the growing demand for health[2] and dedicating their efforts to gaining a competitive advantage by providing medical services tailored to individual characteristics. Diagnostic testing performed as part of the health screening process provides valuable information for risk prediction and diagnosis of diseases, and the results are used as important indicators when judging the examinee's health condition[3]. The study was performed to develop accurate diagnostic models by applying big data analysis and machine learning algorithm models to diagnose diabetes, ventilation, and used health check-ups to determine and predict disease.

2. Literature review

2.1. Concept of health screening

Health screening refers to the patient consultation, physical examination, and other medical examinations conducted at a medical institution for the purpose of assessing the individual's health condition, checking for risk markers, and detecting signs and symptoms of a disease[4]. Health screening can lead to early detection of diseases and enable preventive treatment to prevent disease progression, thereby reducing the risk of mortality as well as related medical costs[5,6]. Population aging and changes in lifestyle and diet are causing a major shift from acute illnesses to chronic illnesses and in leading causes of death. Leading causes of death among adults in recent years are mostly diseases that can be managed by improving one's lifestyle. Chronic illnesses as such can be detected in the early stages through health screenings, and disease progression and aggravation can be prevented by appropriate treatment.

Health screenings are divided into general health screening, oral health screening, and milestone health screening, and the results from the screening tests serve as important indicators of the person's general health. Medical institutions are establishing laboratory information systems (LIS) to systematically manage and analyze health examination results, and ICT needs to be applied to LIS to accurately and quickly analyze large amounts of test results[7]. Past studies have shown that expert systems are useful for healthcare professionals in improving the reliability of diagnostic test results[8]. In this study, clinical pathology examination results from health screenings were analyzed to judge the health screening results.

2.2. Machine learning concept

Machine learning is a branch of Artificial Intelligence that includes the use of big data. The core technology of machine learning is to provide computers or machines with the ability to learn and make predictions based on the data provided using machine learning algorithms[9], which requires an understanding of big data, cloud computing, etc. Machine learning consists of a series of processes in which unprogrammed programs learn patterns from the data provided and then perform appropriate tasks when introduced with new data[10]. In machine learning, ideal learning and prediction models are built using probability and combination theory, mathematical optimization techniques, statistical techniques, algorithms, and computer structures, and it is being developed as a type of convergence technology integrated with methods for researchers to acquire and apply empirical knowledge. Machine learning has recently been defined as a way to predict the future or make decisions using patterns automatically recognized in data as a form of big data analysis techniques[11]. The main goal of machine learning is to build a system that can achieve or exceed human-level capabilities in solving complex tasks or problems to develop predictive models by inputting data to be learned into classification, prediction, and clustering algorithms. To enable computers or machines to learn on their own and develop human-level abilities based on complex data or problems entered[12]. Machine learning techniques are largely divided into supervised learning, unsupervised learning, and reinforced learning[13] (**Figure 1**). Supervised is a method of developing predictive models using classification and prediction algorithms when dependent variables are present, unsupervised learning is a method of developing predictive models using clustering algorithms when dependent variables are not present, and reinforcement learning is a method of supervising self-generated data with high feedback.

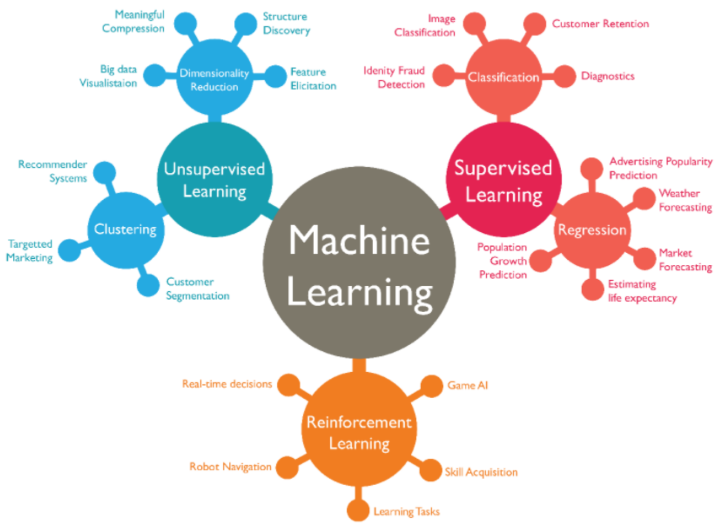

Figure 1. Three classes of machine learning.
Source: (https://www.wordstream.com/blog/ws/2017/07/28/machine-learning-applications).

Machine learning proceeds through the stages of representation, evaluation, optimization, and generalization. Representation means determining the method of processing input values, and this is called a predictive model. Evaluation is a way of determining whether a task has been performed accurately, and optimization is performed to find the condition that best satisfies the evaluation criteria. A predictive model that has passed through the first three steps are said to have completed the learning process, and new data are then entered into the predictive model to predict the results in the generalization step.

2.3. Concept of AI (Artificial Intelligence)

Artificial Intelligence (AI) refers to the intelligence possessed by computers with the ability to learn reason, perceive, and search, which have been considered as abilities unique to humans[14]. The term was first used at the Dartmouth Conference in 1956, when inference and search were key issues in related research[15]. The field of AI went through a period of stagnation until the early 1990s and welcomed a period of revival in the late 1990s with active research on the development of the Internet and machine learning. In the era of the Fourth Industrial Revolution, AI has become a major field of research along with the IoT (Internet of Things), cloud computing, and big data. This type of research has had a significant impact on the industrial sector, and the outcomes have been applied to various fields such as industrial robots, precision medicine, and automatic operation using intelligent learning models[16]. According to a report on the latest technology trends related to Artificial Intelligence from Forbes, AI technology was defined as virtual agents, deep learning platforms, robot process automation, machine learning platforms, biometrics, neural networks, text analysis, natural language generation, text analytics and natural language process, voice recognition, etc.[17].

This study was carried out to provide quality healthcare services and pursue technological development by applying text analysis and a machine learning platform among various technologies related to AI, including the aforementioned technologies, to the healthcare industry.

2.4. Concept of big data analytics

“Big data” means data that are so large and complex data that they cannot be processed by conventional software. Over time, it has become a concept that goes beyond dealing with large amounts of data, and today, it refers to the technological advancement allowing us to not just use big data and but also leverage the opportunities provided by data[18]. With big data becoming a key word in technology, data analytics and utilization is regarded as a crucial field. Most traditional data analysis techniques involve storing data collected from a single computer in a database, performing analysis, or performing operations based on stored data. With the advent of big data, a new technology and method for processing a wide range of data, Facebook, YouTube, Twitter, and blog services are conducting in-depth analyses of customer data based on access logs and transaction logs. Research on these analytical techniques has been conducted extensively across all areas. Big data technology can be divided into analysis and infrastructure. Many of the applied analytical techniques are already widely used in statistics, computer science, machine learning, data mining, and so on, and can be applied properly because they are continuously upgraded and optimized for big data processing. Big data analysis methods include data mining, machine learning, natural language processing, and pattern recognition used in statistics. Recently, text mining, opinion mining, social network analysis, and cluster analysis have attracted attention due to the accumulation of data and the increase in unstructured data such as social media due to the spread of smartphones. Big data analysis systems will probably be applied most actively among various types of Artificial Intelligence technologies. A method of expressing the professional or empirical knowledge of experts in a particular field within a computer to solve problems without the help of experts or to help them solve problems[19]. The big data analysis procedure can be divided into pre-processing and post-processing steps: pre-processing includes data filtering, data type conversion, and refining and post-processing includes data transformation, integration, and reduction. In the healthcare sector, in particular, medical institutions use big data for the purpose of providing high-quality medical services to patients. It is used for patients, the medical staff, patient management, and business management support. With the transition from the era of information technology (IT) characterized by creation of information and knowledge using data to the era of data technology (DT) with the development and evaluation of services, the field of big data application in healthcare is also expanding[20]. Along with these changes in the utilization of big data in healthcare, there has been a growing demand for high-quality big data in healthcare, and it has been noted that there is a need to achieve qualitative improvement to ensure efficient utilization. The World Health Organization (WHO) published a guideline for improving the quality of health and medical data in 2003[21]. It provides instructions to healthcare workers and healthcare information managers of various ranks in the organization and offers essential information for all areas associated with collecting and generating data. The guideline requires that data quality measures be developed in consideration of the importance, accuracy, reliability, validity, completeness, currency and timeliness, legibility, accessibility, and confidentiality of data.

In this study, a platform called Hello Data, where users can select the tools necessary for creating a project and processing data and create their own templates, was used to ensure an efficient data processing service based on fast and accurate processing and annotation using a pre-annotation feature.

2.5. Tests to diagnose diabetes mellitus

Diabetes mellitus is a disorder in which glucose builds in the bloodstream instead of being absorbed by cells, which need it to use as their energy source, due to insulin resistance or deficiency, and ends up being excreted in the urine. The three main symptoms of diabetes that are useful in self-diagnosis are polyuria (increased frequency of urination), polydipsia (increased thirst), and polyphagia (increased hunger). When these symptoms appear, it means that the insulin secretion function of the pancreas has significantly deteriorated already, making it necessary for the individual to manage their blood pressure and obesity through regular health screenings. For the diagnosis of diabetes, urine sugar, blood sugar, and glucose tolerance test results are used along with the information on family history and clinical symptoms, and diagnosis is made by also considering the glycated hemoglobin test and C-peptide test.

Diabetes can be diagnosed by a urine glucose test, which is one of the most commonly performed tests. Measuring the urine glucose level is arguably the simplest and fastest test method for diabetes diagnosis, but it has relatively poor accuracy. This is because there may be dietary, neurological, or other reasons for glycosuria aside from diabetes. In addition, since glucose in the urine is detected only when the concentration is at least 180 mg/dL, making it difficult to diagnose diabetes in the early stages. Urine tests can be performed to check the specific gravity of urine, urine ketone bodies, and so on, in addition to measuring the glucose concentration. Secondly, a blood glucose test may be performed, and it can even be self-administered. Normally, blood sugar rises temporarily after a meal, but excess sugar is stored in the liver as glycogen or stored as fat by the body to keep the blood sugar constant. On the other hand, people with diabetes have a high fasting blood sugar level, so it is usually tested on after fasting for 12 to 18 hours for accurate results. A fasting blood glucose level falls within the range of 70 to 100 mg/dL is normal, but it may be 140 to 300 mg/dL or even higher in people with diabetes. Thirdly, a glucose tolerance test may be performed to determine a person's ability to handle a glucose load if a conclusive diagnosis cannot be made based on a blood glucose test. It involves injection or oral administration of 75g of glucose after fasting, and blood is collected five times immediately after the administration and at 30, 60, 90, and 120 min afterwards to measure the blood sugar level and plot a blood sugar concentration curve. In nondiabetic individuals, blood sugar peaks between 30 to 60 min and declines gradually afterwards before returning to normal levels after two hours. However, if the examinee has diabetes, blood sugar will continue to rise without dropping to normal levels even after two hours. Fourthly, a glycosylated hemoglobin test may be performed, and this is based on the fact that high blood glucose levels lead to an increase in glycated hemoglobin. It involves measuring the concentration of glycated hemoglobin to evaluate the mean blood glucose level over a period of two to three months. Glycated hemoglobin is generally used as an indicator that serves useful in diabetes management because it can be used to determine the long-term average level of blood sugar. When blood sugar rises, glucose in the blood non-enzymatically bonds with hemoglobin, and the amount of glycated hemoglobin, with glucose attached to the end of the molecular structure of HBA1C, Increases. In normal, healthy Individuals, glycated hemoglobin accounts for about 6 to 7.5% of the total hemoglobin, but in diabetic patients, it is nearly double that amount at around 11%. This test can be administered regardless of whether the person has eaten or not, unlike the fasting blood sugar test. Plus, the amount of glycated hemoglobin does not fluctuate as easily as blood sugar, so it is used as an indicator to determine whether blood sugar has been controlled well in recent months. IFG (Impaired fasting glucose) is a condition in which the fasting (no caloric intake for at least 8 hours) blood sugar is in the range of 100 to 120 ml/dL, which is lower than the reference level for diabetes diagnosis, but higher than normal. That is, the blood sugar level is higher than normal due to an impaired function related to insulin secretion. IFG patients is in a prediabetes stage and may potentially develop diabetes if they continue their current lifestyle. On the other hand, IGT (impaired glucose tolerance) is diagnosed when the blood glucose level is between 140 and 199 mg/dL two hours after a meal. Those with IGT tend to have family history of diabetes and suffer from obesity, hypertension, and/or hyperlipidemia, and they may even have IFG. Lastly, a C-peptide test may be performed for diabetes diagnosis. C-peptide is secreted in a 1:1 ratio when insulin is synthesized from proinsulin, and by continually measuring the C-peptide level, it is possible to predict the insulin secretion time and amount after sugar intake. This is a test method that helps determine whether an appropriate amount of insulin is being administered for diabetes management.

Diabetes management aims to regulate blood sugar to near-normal levels so as to prevent possible complications, such as diabetic retinopathy, gangrene, nerve disorders, and blood vessel damage. Blood pressure and obesity management is key to leading a normal life, and efforts must be made by the patient to lower their blood sugar by changing their diet and exercising.

2.6. Tests to diagnose gout

Gout is a chronic systemic condition in which recurrent paroxysmal inflammation occurs in the joints and surrounding tissues due to phagocytosis of urate crystals by leukocytes, as urate crystals accumulate in the body due to abnormal purine metabolism and impaired renal excretion of uric acid[6,7]. Gout is known to be associated metabolic syndrome as well as arthritis and cause severe pain[8]. For a conclusive diagnosis of gout, the affected joint or soft tissue gets punctured for proof of presence of urate crystals in the synovial fluid, tissue, or gout nodule (tophi). When uric acid gout crystals are viewed under a polarized light microscope, they will appear as strongly negative birefringent, needle-shaped crystals. This is the gold standard for diagnosing gout[22]. If it is difficult to collect synovial fluid or tissue, a definite diagnosis cannot be made, but in this case, gout can be diagnosed based on the gout classification criteria published jointly by ACR and EULAR in 2015[23].

According to the gout classification criteria, gout is diagnosed if the sum of the scores from the clinical criteria, laboratory criteria, and imaging criteria is 8 points or higher (out of 23 points). In the clinical criteria, a score between 1 and 3 points is assigned according to the location of onset, symptoms, and duration. In the laboratory criteria, a score is given based on the serum uric acid concentration: 2 points if it is between 6 and 7.9 mg/dL, 3 points for 8 to 9.9 mg/dL, and 4 points for 10 mg/dL or higher. If the serum uric acid concentration is less than 4 mg/dL, 2 points are deducted, and if the synovial fluid test is performed and no urate crystals are detected, 2 points are deducted as well. As for the imaging criteria, 4 points are added if there is imaging evidence of urate deposition such as a double contour sign (DCS) indicating a buildup of uric acid on articular cartilage, which is a characteristic of gout, in articular ultrasonography or detection of urate deposition in joint, periarticular bursa, ligament, muscles, etc., and 4 points are added if there is imaging evidence of gout-related joint damage in conventional X-ray.

According to these criteria, it is important for the doctor to listen closely about the patient's medical history related to the clinical symptoms, and for a definitive diagnosis of gout, a physical examination, serum uric acid test, articular ultrasonography, DECT, and conventional radiography are necessary.

3. Research methodology

3.1. Objective of the research

This study was carried out for the purpose of developing an assessment and prediction module for diabetes, gout by analyzing the diagnostic test results of health screening examinees using machine learning technology based on big data analysis to improve the quality of healthcare services for increased quality of life and life expectancy. The research procedure for the module development consisted of three stages, processing of high-quality data, establishment of an AI model, and business application, as shown in **Figure 2**.

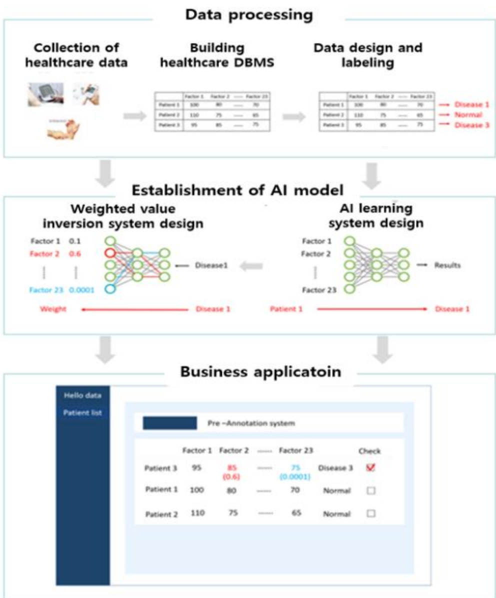

Figure 2. Research procedure.

3.2. Method of the research

Research procedures were established for research design, and high-quality data collection and supplementation, analysis data preprocessing, predictive model development, and verification were carried out according to the research procedures. To develop a predictive model, we used a random forest with 23 diagnostic test factors and 2 demographic factors. Predictive model performance was measured and analyzed based on TP speed, FP speed, precision, accuracy, and ROC area metrics. Among the predictive model performance indicators, accuracy and ROC area were used to determine the classification accuracy and the validity of predictions, respectively. The first step in the research procedure was pre-processing. The current healthcare systems predict disease onset based on a single indicator, but in this study, where there were plans to make predictions based on a comprehensive evaluation of the results of 23 diagnostic tests, the data from each test were normalized. Al, given the fact that the diagnostic test results obtained by different medical institutions tend to vary to a certain extent, feature extraction was performed through dimension reduction to determine the type of data to be used as the key data or incidental data for each disease. The second step was data generation. The data processed in the data preprocessing step were entered into the data processing platform, Hello Data, to edit the two disease profiles. Here, data were generated to meet the target by using the pre-annotation and data analysis features on Hello Data. The third step was the data augmentation process. Since large amounts of data are necessary for learning of algorithms, a distribution function is set based on the data obtained during the data generation process. The feature extraction results from the data preprocessing process were referenced to augment the data so that the AI algorithm would not be affected even if there were no results from secondary diagnostic tests. Therefore, sufficient data was obtained for fast learning of algorithms by forming a distribution function of the results from each of the 23 diagnostic tests and generating tens of thousands of virtual data to maintain the correlation and independence between the distribution functions. The fourth step was the AI algorithm development and learning process. By finding and training a suitable AI model for the type of data that would be used and repeating the entire process while providing feedback on the results, the data and system development to be carried out in this study was completed.

3.3. Analysis of diabetes prediction

The positive cases of diabetes accounted for 21% of the data used in the predictive model. This was deemed an appropriate level, as the data were to be used for modeling at the design level. First, in order to determine the accuracy of diabetes prediction and the factors affecting diabetes prediction, a random forest was used to assess the impact of 25 factors on diabetes, and the results are shown in **Figure 3**. The accuracy of predicting diabetes was found to be 0.91, and an analysis of the importance of features showed that HBA1C, GLUCOSE, and LDL among the 23 tests had the highest impact in predicting diabetes.

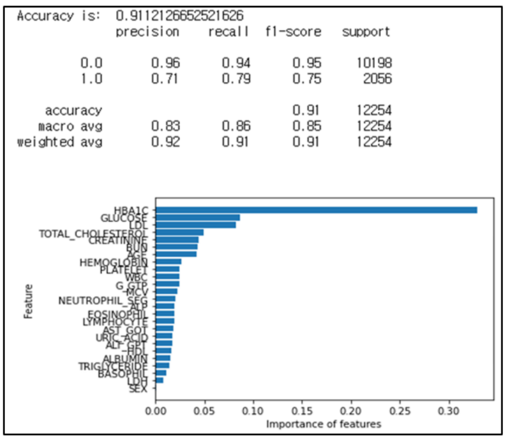

Figure 3. Random forest analysis (diabetes prediction).

Second, for accuracy analysis, the loss and accuracy were analyzed, and the result was 0.86. A graph of the results is shown in **Figure 4**.

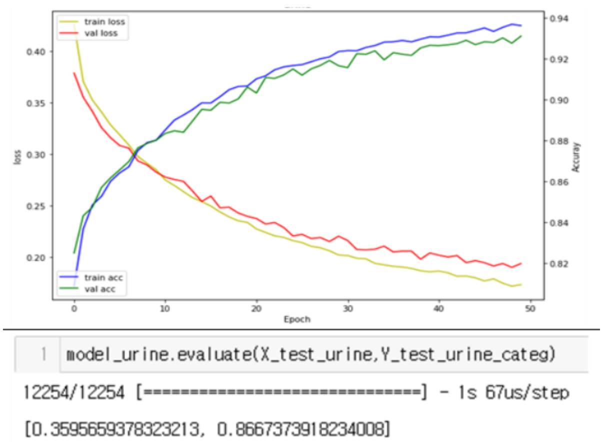

Figure 4. Loss and accuracy analysis results (diabetes prediction).

An analysis of the ROC curve, a method of evaluating the performance of binary classification algorithms, shows that the area under the curve (AUC) was 0.92, indicating high predictive power. A graph of the results is shown in **Figure 5**.

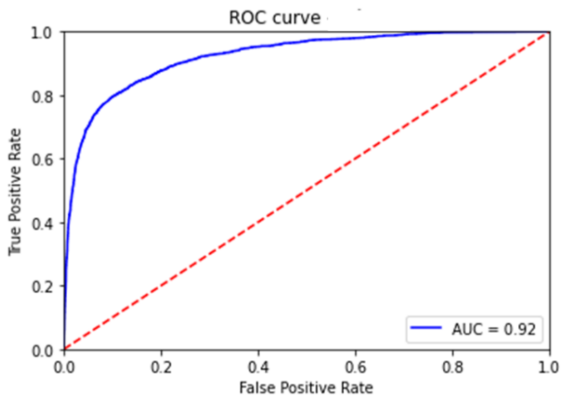

Figure 5. ROC curve (diabetes prediction).

The confusion matrix analysis results for diabetes prediction are shown in **Figure 6**. The TP rate and FP rate were 0.88 and 0.12, respectively, and precision and recall were 0.88 and 0.82, respectively. These were deemed satisfactory.

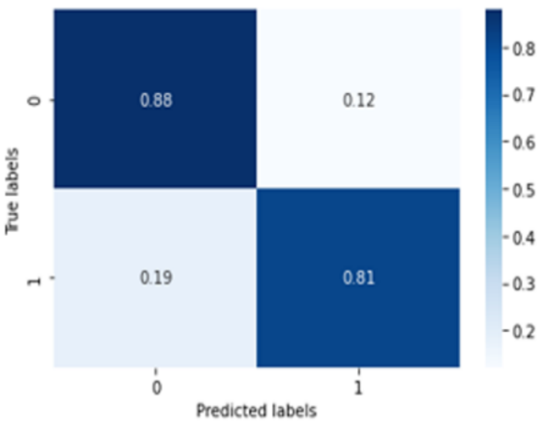

Figure 6. Confusion matrix (diabetes prediction).

3.4. Analysis of gout prediction

Diabetes-positive patients accounted for 28% of the data used in the predictive model. First, the effect of 25 factors on gout was evaluated using a random forest to determine the accuracy of diabetes prediction and the factors affecting gout prediction, and the results are shown in (**Figure 7**). The accuracy of predicting gout was found to be 0.98, and an analysis of the importance of features showed that CREATININE, URIC_ACID, and G_GTP among the 23 tests had the highest impact in predicting gout.

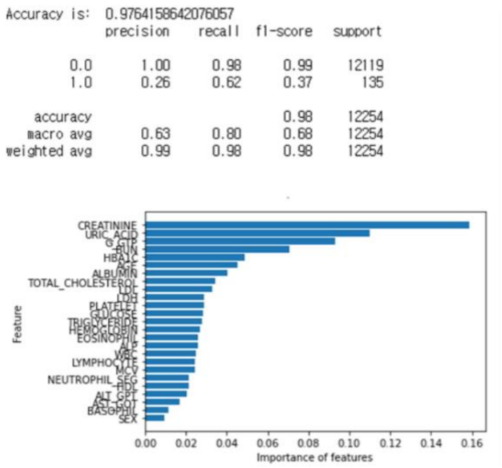

Figure 7. Random forest analysis results (gout prediction).

Second, for accuracy analysis, the loss and accuracy were analyzed, and the result was 0.96. A graph of the results is shown in **Figure 8**.

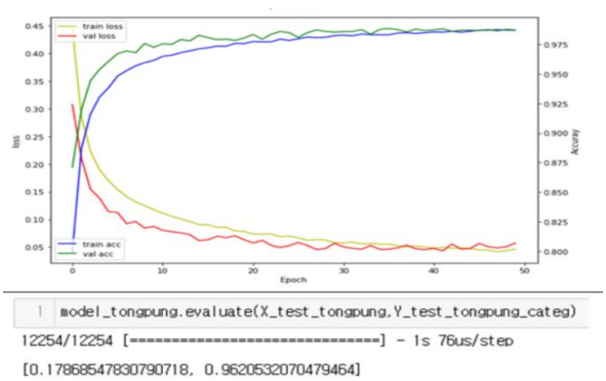

Figure 8. Loss and accuracy analysis results (gout prediction).

An analysis of the ROC curve, a method of evaluating the performance of binary classification algorithms, shows that the AUC was 0.99, indicating high predictive power. A graph of the results is shown in **Figure 9**.

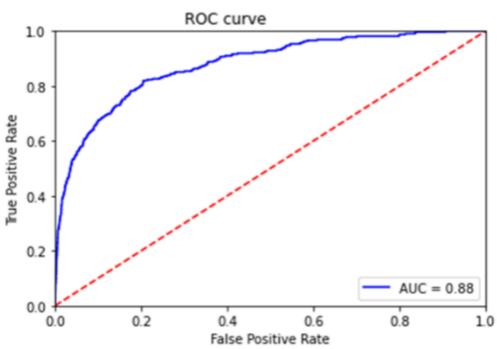

Figure 9. ROC curve (gout prediction).

The confusion matrix analysis results for gout prediction are shown in **Figure 10** The TP rate and FP rate were 0.98 and 0.024, respectively, and precision and recall were 0.97 and 0.63, respectively. As such, precision was good, but recall was low, based on which it was determined that the accuracy of the predictive model in judging abnormality was low.

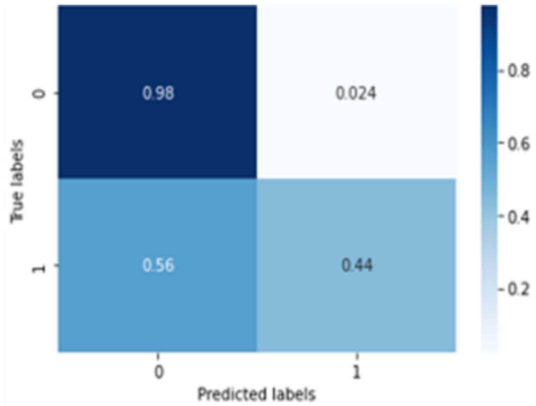

Figure 10. Confusion matrix (gout prediction).

4. Result and discussion

The indicators that determine the accuracy of disease diagnosis and prediction models were accuracy, AUC, TP speed, FP speed, precision, and recall. When determining the performance of the mode, a separate test set was applied to the modeling dataset, and the measurements were supplemented by calculating the standard deviation (SD) of the error through 10-fold cross-validation to determine the likelihood of discrepancies depending on the sample used. A random forest was applied to process the data from the 23 tests and on sex and age and the pre-annotation and data analysis functions of Hello Data, a data processing tool, were used to obtain sufficient data for the AI to learn the algorithms for prediction of three diseases: diabetes, gout, and anemia (see **Table 1**). The disease prediction model was developed by repeating the process of finding a suitable AI model, training it, and giving feedback on the prediction results, and the model developed may be used as a baseline model to develop the disease prediction model meeting the target requirements in the future.

Table 1. Results of analyzing the disease prediction model.

Disease	Accuracy	AUC	Precision	Recall
Diabetes	0.91	0.92	0.88	0.82
Gout	0.98	0.88	0.97	0.63

According to similar studies related to diabetes, if a prediction model uses Naive Bayes, Decision Tree, and deep learning and the accuracy is more than 90%, it can be used as a model for a commercial platform. Although the objective value of the world's highest level of accuracy is not known, the accuracy of the disease prediction model developed in this study in predicting diabetes was found to be 91%, which means it has utility as a prediction model on a commercial platform. However, gout was predicted with an accuracy of 85%, which did not meet the standard for application to commercial platforms. The results of the ROC curve analysis showed that the AUC for prediction of three diseases ranged from 0.88 to 0.92, indicating that significantly high predictive power for all two diseases. Precision was found to be in the range of 0.88 to 0.97, indicating high reliability. Recall was satisfactory for diabetes prediction at 0.82, but low for gout prediction at 0.63. As a result of analyzing the results of the confusion matrix to determine the cause of the low recall for gout prediction, the true negative rate was found to be 98% for gout, and the false negative rate was 56% for gout. This meant the error in judging abnormality was high.

5. Conclusion and future work

This study was carried out to develop modules for predicting diabetes, gout based on health screening results through the application of big data-based machine learning. For this purpose, the Hello Data platform was used, and learning data and test data were used separately to develop the disease prediction model and to evaluate its performance. Random forest was used to develop the predictive model, and accuracy, area under the receiver operating characteristic (ROC) curve, confusion matrix, precision, and recall were used as indicators for evaluating the performance of the predictive model. The results were as follows:

First, an analysis of the disease prediction model showed that the accuracy of predicting diabetes, gout, and anemia was at least 90%, based on which it was judged that it would be appropriate for the commercial platform. The diagnostic tests that were found to have a high level of importance in disease prediction were HBA1C, GLUCOSE, and LDL for diabetes, CREATININE, URIC_ACID, and G_GTP for gout, and HEMOGLOBIN, CREATININE, and BUN for anemia. Just as how the results from various diagnostic tests are used in combination to diagnose a disease in the actual healthcare setting, the results of multiple tests were used to predict diseases in this study.

Second, ROC curve analysis was carried to analyze the utility of the diagnostic module by evaluating the performance of the predictive model developed as part of the learning module. The area under the curve (AUC) was found to be 0.92 for diabetes, 0.88 for gout, and 0.90 for anemia, indicating high utility. Third, an analysis of the confusion matrix for an evaluation of the prediction performance showed that precision, the ratio of true positives to the total number of elements labelled as belonging to the positive class, was 0.88 for diabetes, 0.97 for gout, and 0.94 for anemia, which were all high. Based on this, it was deemed that the model was suitable for predicting the absence of the target disease. As for recall, the ratio of true positives to the total number of elements that actually belong to the positive class (i.e., the sum of true positives and false negatives), was 0.82 for diabetes, 0.63 for gout, and 0.72 for anemia. As such, recall in diabetes prediction was satisfactory, but it was low for gout and diabetes. As a result of analyzing the causes of the low recall in predicting gout and anemia, the true positive (TP) rate was found to be 0.98 for gout and 0.94 for anemia, which were both very high, while the false positive (FP) rate was 0.024 and 0.053, respectively, which were low. Based on this, it was found that positive results were predicted correctly with strong discriminating power. On the other hand, the true negative (TN) rate was low at 0.44 for gout and 0.63 for anemia, and the false negative (FN) rate was high at 0.56 for gout and 0.37 for anemia. As such, it was found that the disease prediction model recorded low TP rates and high FP rates, based on which the discriminating power was deemed insignificant. Considering the high FP rates in predicting gout and anemia, it is deemed necessary to conduct research to lower the FP rates through supplementation of data and repeated learning.

As for the implications of this study, it is believed it will be possible to expand upon the learning model developed in this study to automatically predict other diseases. This will lead to the development of a new healthcare service model that can improve the health screening system and address related issues, thereby enhancing the reliability and efficiency of health screenings. Also, considering that there is strong correlation among various chronic diseases, it will be possible to diagnose multiple chronic diseases using the disease prediction model by simply using different data.

Author contributions

Conceptualization, JYP and JWO; methodology, JYP and JWO; software, JYP and JWO; validation, JYP and JWO; formal analysis, JYP and JWO; investigation, JYP and JWO; resources, JYP and JWO; data curation, JYP and JWO; writing—original draft preparation, JYP and JWO; writing—review and editing, JYP and JWO; visualization, JYP and JWO; supervision, JYP and JWO; project administration, JYP and JWO; funding acquisition, JYP and JWO. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

This research was supported by “Regional Innovation Strategy (RIS)” through the National Research Foundation of Korea (NRF) by the Ministry of Education (MOE) in 2023 (2022RIS-005).

This research was supported by Kyungdong University Research Fund, 2022.

Conflict of interest

The authors declare no conflict of interest.

References

1. Kim HS. A Study on the Efficient Policy of Health Examination based on Comparing Private Health Sector with Public Health Sector [PhD thesis]. Kyunghee University; 2010.
2. Min WG. New Criteria for the Quantitative Evaluation of Quality Control Results [PhD thesis]. Seoul National University; 1992.
3. Kim YT, Chae BS, Hwang BJ. The effect of physical environments in the comprehensive health examination center on medical service value, satisfaction and switching barrier (Korean). Journal of Service Research and Studies 2019; 9(4): 63–80. doi: 10.18807/jsrs.2019.9.4.063
4. Bartels PH, Thompson D, Wedber JE. Diagnostic and prognositc decision support systems. Pathologica 1995; 87(3): 221–236.
5. Swets JA. Measuring the accuracy of diagnostic systems. Science 1988; 240(4857): 1285–1293. doi: 10.1126/science.3287615
6. Singh JA, Strand V. Gout is associated with more comorbidities, poorer health-related quality of life and higher healthcare utilisation in US veterans. Annals of the Rheumatic Diseases 2008; 67(9): 1310–1316. doi: 10.1136/ard.2007.081604
7. Lee SJ, Hirsch JD, Terkeltaub R, et al. Perceptions of disease and health-related quality of life among patients with gout. Rheumatology 2009; 48(5): 582–586. doi: 10.1093/rheumatology/kep047
8. Choi HK, Ford ES, Li C, Curhan G. Prevalence of the metabolic syndrome in patients with gout: The third national health and nutrition examination survey. Arthritis & Rheumatism 2007; 57(1): 109–115. doi: 10.1002/art.22466
9. Oh JW, Kang JK. Exploration of disease occurrence prediction model based on clinical pathology results using machine learning. In: Tripathy HK, Mishra S, Mallick PK, et al. (editors). Technical Advancements of Machine Learning in Healthcare. Springer, Singapore; 2021. pp. 377–388.
10. Krizhevsky A, Sutskever I, Hilton GE. ImageNet classification with deep convolutional neural networks. Communications of the ACM 2017; 60(6): 84–90. doi: 10.1145/3065386
11. Murphy KP. Machine Learning: A Probabilistic Perspective. The MIT Press; 2015.
12. Buyya R, Calheiros RN, Dastjerdi AV. Big Data: Principles and Paradigms, 1st ed. Morgan Kaufmann; 2016.
13. Géron A. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, 2nd ed. O’Reilly Media; 2019.
14. Kim SR. The current situation and the suggestions of AI & Law research in legal reasoning (Korean). IT and Law Review 2011; 5: 319–346.
15. Priyopradono B, Manongga D, Utomo WH. Spatial social network analysis: Program Pengembangan Usaha Agribisnis Perdesaan (PUAP) or an exertion development program in supporting the region revitalization development. Social Networking 2013; 2(2): 63–76. doi: 10.4236/sn.2013.22008
16. Lu H, Li Y, Chen M, et al. Brain intelligence: Go beyond Artificial Intelligence. Mobile Networks and Applications 2017; 23(2): 368–375. doi: 10.1007/s11036-017-0932-8
17. Press G. Top 10 hot Artificial Intelligence (AI) technologies. Available online: https://www.forbes.com/sites/gilpress/2017/01/23/top-10-hot-artificial-intelligence-ai-technologies/?sh=68ea0f121928 (accessed on 3 August 2023).
18. Lee HT, Lee SW, Cho JW, Cho IS. Analysis of feature importance of ship’s berthing velocity using classification algorithms of machine learning. Journal of the Korean Society of Marine Environment & Safety 2020; 26(2): 139–148. doi: 10.7837/kosomes.2020.26.2.139
19. Waterman DA. A Guide to Expert Systems. Addison-Wesley Publishing Company; 1986.
20. Cho IS. Assessing the quality of structured data entry for the secondary use of electronic medical records. Journal of Korean Society of Medical Informatics 2009; 15(4): 423–431. doi: 10.4258/jksmi.2009.15.4.423
21. World Health Organization. Improving Data Quality: A Guide for Developing Countries. WHO Regional Office for the Western Pacific; 2003.

22. Malik A, Schumacher HR, Dinnella JE, Clayburne GM. Clinical diagnostic criteria for gout: Comparison with the gold standard of synovial fluid crystal analysis. Journal of Clinical Rheumatology 2009; 15(1): 22–24. doi: 10.1097/RHU.0b013e3181945b79
23. Neogi T, Jansen TL, Dalbeth N, et al. 2015 gout classification criteria: An American college of rheumatology/European league against rheumatism collaborative initiative. Arthritis & Rheumatology 2015; 67(10): 2557–2568. doi: 10.1136/annrheumdis-2015-208237