Int. j. adv. multidisc. res. stud. 2026; 6(2):1562-1566

Received: 27-02-2026
Accepted: 07-04-2026

International Journal of Advanced Multidisciplinary Research and Studies

ISSN: 2583-049X

Prediction Performance of Liver Disease based on GA with ML

[1] **Dr. Rachhpal Singh,** [2] **Dr. Parvinder Kaur,** [3] **Nitish Sharma**
[1] PG, Department of Computer Science and Applications, Khalsa College, Amritsar, India
[2, 3] I.K. Gujral Punjab Technical University Campus, Batala, India

DOI: https://doi.org/10.62225/2583049X.2026.6.2.6138
Corresponding Author: **Dr. Rachhpal Singh**

Abstract
Complex optimization issues are increasingly being solved by combining machine learning (ML) with genetic algorithms (GA). In high-dimensional, non-convex spaces, this hybrid method (ML-GA) is very helpful for locating global optima. The LD (Liver Disease) datasets are converted into an image-like input for the ML architecture by using both as information gains for feature selection optimization. By using a GA for better prediction performance in LD detection, this will get around the drawbacks of conventional hyper-parameter optimization methods. The sophisticated GA-based machine learning model significantly improved accuracy by outperforming conventional techniques. For both binary and multiclass LD prediction tasks, the optimized model produced a promising accuracy range using machine learning algorithms, including Naïve Bayes (NB), Support Vector Machine (SVM), Decision Tree (DT), Logistic Regression (LR) and Random Forest (RF).

Keywords: Genetic Algorithm, Liver Disease, Artificial Intelligence, Machine Learning, Deep Learning, Healthcare System, Optimization Techniques

1. Introduction

Excess fat (of liver weight) accumulates in liver cells to cause fatty liver disease, which frequently has no symptoms but may result in inflammation, cirrhosis, or cancer [1]. Fatty liver leads to liver disease (LD), one of the World Health Organization's top causes of death. About 10.2 million lives are lost to LD each year, highlighting the significance of early detection and individualized treatment to lower patient risk [2]. Artificial intelligence (AI), machine learning, and deep learning are increasingly being used in healthcare; deep learning in particular has the potential to improve the prognosis of LD and personalize patient care [3]. Even though these technologies are essential for improving prediction accuracy [4], it is still difficult to forecast LD [5]. By creating models that make use of known risk factors like age, sex, blood pressure, cholesterol levels, smoking habits, and family medical history, machine learning has demonstrated encouraging results in the prediction of fatty liver disease. These forecasts have made use of important methods such as support vector machine (SVM), k-nearest neighbor, XGBoost, naive Bayes, and multilayer perceptron [6].

A key factor in improving the precision, ease of use, and effectiveness of these models is feature selection, especially through information gain (IG). This results in more precise, comprehensible models that are ideal for real-world use in a variety of contexts [7]. In order to optimize the prediction process, IG plays a crucial role in determining the most important characteristics for LD prediction [8]. A key component of machine learning is hyper-parameter optimization, which seeks to balance variation in order to increase model accuracy and avoid over-fitting. This procedure entails establishing particular model configurations that are chosen to improve network performance but are not based on data. Because the values of a model's hyper-parameters have a major impact on its effectiveness [9], common techniques like grid search [10], randomized search [11], and Bayesian optimization [12] are often used. In order to maximize performance indices, the wide variety of hyper-parameter combinations poses a significant challenge [13] and is tackled as a multi-objective problem [14]. By investigating ideal configurations, meta-heuristic techniques like the Ant Colony System [15] and Genetic Algorithms (GAs) [16] solve this problem.

When a Convolutional Neural Network (CNN) is used for feature engineering, the network itself automatically extracts and represents pertinent features from raw input. By automatically identifying discriminative features from raw data, the CNN's convolutional layers are specifically engineered to learn and extract features hierarchically, hence minimizing the need for manual feature engineering [17]. With this method, human-designed characteristics are no longer necessary, and the network can recognize crucial patterns in the data on its own. Depending on enough data for their design and training, CNNs provide excellent accuracy and flexibility and enable improved categorization in a variety of areas [18]. For optimal performance, figuring out the hyper-parameters is essential. In applications including crop pest picture classification [19], croup cough classification [20], and pattern recognition [21], GA-based CNNs have shown satisfactory performance. Additionally, they could enhance the results of LD prediction. Compared to conventional techniques, using a GA-based CNN for hyper-parameter optimization has a number of benefits. In order to avoid local minima and find the best hyper-parameters, GAs are especially good at striking a balance between exploration and exploitation. They are computationally superior to exhaustive techniques like grid search because of this feature. Furthermore, GAs are quite flexible and can handle the intricate, high-dimensional search spaces that are characteristic of deep learning models. Large models may be efficiently optimized thanks to their scalability, and the crossover and mutation procedures help avoid over-fitting. Furthermore, GAs' parallelizability increases their efficiency by enabling the evaluation of several solutions at once. All things considered, this method offers a reliable and adaptable way to optimize LD prediction models.

Table-based public datasets for LD prediction include the Cleveland dataset from the UCI Machine-Learning Repository. Nevertheless, CNNs face difficulties when handling tabular data because they are mostly made for picture data. By converting tabular data into images [22], spatial information is preserved, allowing CNNs to efficiently identify spatial patterns and connections [23]. This technique enhances the model's capacity to identify complex linkages and patterns that may not be entirely apparent when using tabular data directly with a CNN. In order to maximize the precision of early LD detection, this work presented a GA-based CNN feature engineering technique. This study's primary contributions are:

1. The use of IG for critical feature selection.
2. The application of GA for CNN hyper-parameter tuning to enhance early LD detection prediction performance.
3. The conversion of tabular data into image format to better capture spatial correlations and increase LD prediction accuracy.

2. Related Work

On the SO online Q&A page, no one SLR study has been conducted on the use of machine learning algorithms or approaches to determine software needs. Although there are several surveys on the SO [24], none of them currently have anything to do with machine learning algorithms or techniques for identifying or detecting software needs on the SO online Q&A website. Ahmad *et al.* [25] conducted an ad hoc literature review on the SO in 2018, which mostly focused on the software development life cycle from the website's launch to 2016.

In a similar vein, Baltadzhieva and Chrupała's work [26] carefully examined and evaluated the quality of several questions that were posted on various community question answering (CQA) websites, such as SO. Other studies are available [27] that highlight machine learning algorithms used to discover, gather, and classify nonfunctional requirements in software documentation. An SLR on the literature on automated requirements elicitation was carried out in 2013 by Meth *et al.* [28].

SLR on ML algorithms for detecting and categorizing NFRs was carried out by Binkhonain and Zaho [27]. They have chosen 24 published primary studies. The main conclusions of their research showed that ML techniques may detect and categorize NFRs, but they still face numerous difficulties that require further study. A survey on requirements engineering and machine learning techniques was recently presented by Iqbal *et al.* [29]. They gave an overview of how various requirements engineering tasks are being aided by ML algorithms.

Additionally, a number of surveys, SLRs, and systematic mapping studies have been conducted in various fields, such as sentiment analysis of scientific citations [30], data preprocessing techniques for the problem of class imbalance [31], software development effort estimation models based on ML algorithms or techniques [32], usability in agile software development [33], and requirements prioritization [34].

3. Methodology

A branch of computer science called machine learning (ML) makes use of computer algorithms to find patterns in massive amounts of data and help forecast different outcomes based on data [35]. In several fields, machine learning approaches have become a promising tool for forecasting and making decisions [36]. ML has also proven crucial in medical decision making because of the accessibility of clinical data [37]. Creating a machine learning model would be a useful tool for diagnosing illnesses and making efficient clinical decisions in real time. Additionally, by correctly categorizing patients with significant multiple risk factors earlier, it would enable the optimization of hospital resources [38].

The Cleveland dataset from the UCI Machine Learning Repository was used in this investigation. With an emphasis on the objective feature that signifies the existence of liver illness, this database has over 80 features. Research has often focused on a subset of 17 characteristics from over 400 patient records. The dataset is mostly utilized for binary classification in order to differentiate between liver disease instances. As seen in Figure 1, CNN is improved by a GA and created especially for the early detection of LD. The methodology was organized as follows: first, an LD dataset was obtained and pre-processed to deal with missing values. The strongest predictive characteristics were then found using an IG approach and transformed into an image-like format appropriate for CNN analysis. A GA was used to further improve the CNN's feature extraction capabilities by optimizing its hyper-parameters. Several machine learning classifiers, such as naive Bayes, SVM, decision tree, and logistic regression, were then trained using the optimized features. This all-encompassing strategy, which combines data preparation, feature engineering, and machine learning, significantly increased the accuracy of LD prediction, according to a thorough evaluation using metrics like accuracy, precision, recall, and F-measure.

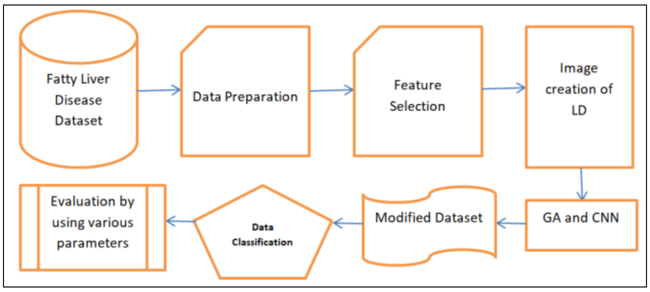

Fig 1: Proposed GA with CNN for LD

IG was used to find the most predictive features for feature selection, as shown in Figure 2(a). This required a number of procedures, including computing IG, sorting the features, calculating entropy, and figuring out the conditional entropy for each feature. In order to ensure that the model emphasizes the important predictors of LD, this step focuses on choosing the most informative features. The next task is to convert the tabular LD dataset into a three-dimensional, image-like structure in order to adapt it for CNN application. This transformation process is shown in Figure 2(b), where the dimensions of the dataset are defined, the number of channels is specified based on grouped categories from the chosen features, and they are configured into a 3 × 4 × 1 input shape for CNN analysis.

The input layer of the GA-based CNN in Figure 2(c) is made to support a 3 × 4 × 1 representation of the dataset's features. To increase training speed and stability, the convolutional layer performs many convolution operations, each followed by batch normalization. The CNN's hyper-parameters are optimized by a GA. It uses crossover and mutation, creates possible solutions (sets of hyper-parameters), and assesses their fitness, including accuracy. In the end, an optimal set is reached by selecting the best solutions for successive generations. The streamlined feature set produced by this optimized CNN improves the predictive power of conventional machine learning models for LD. Metrics including accuracy, precision, recall, F1-score, specificity, G-mean, and p-value are used to thoroughly assess the performance.

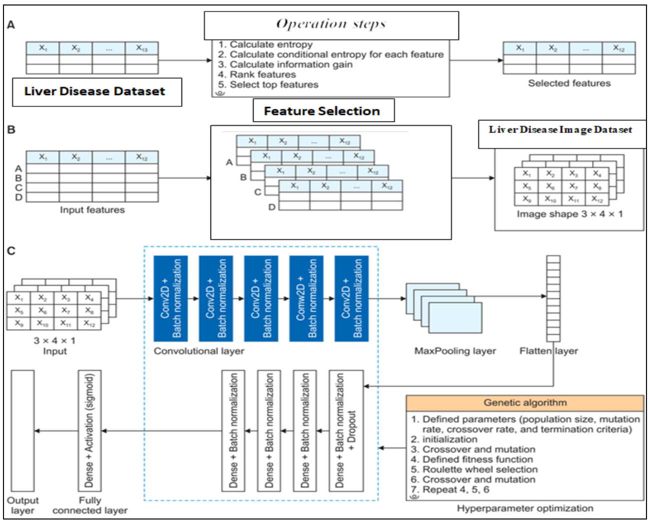

Fig 2: Feature selection Process (a), Data conversion to Image data (b), GA with ML implementation (c)

In order to provide a thorough picture of the model's capabilities, an evaluation was carried out to evaluate the performance of LD prediction using a number of measures beyond accuracy. Recall, precision, F1-score, specificity, and G-mean were all analyzed. To further illustrate each metric's statistical importance, we provide the p-value. A more comprehensive grasp of the model's advantages and disadvantages across several performance characteristics is provided by this enhanced evaluation.

4. Results and Discussion

To guarantee data completeness and integrity, the LD dataset was carefully checked for missing values. Inspection revealed that neither the binary nor the multiclass datasets in the LD dataset had any missing values. The pre-processing step is made simpler by the lack of missing values, which eliminates the need for imputation techniques and lets us move straight to feature selection and modeling. Cholesterol, maximal liver fatty rate, chest pain, ST depression, thalassemia, visible vessels, exercise angina, age, liver rate slope, resting blood pressure, sex, and Cretonne findings are the relevant features in the binary dataset along with their IG values. After the LD dataset was converted into an image representation, 303 samples made up the CNN's input shape. Think about the characteristics of a single LD sample, such as age, blood pressure, heart rate, and cholesterol. These characteristics are transferred onto three different color channels, such as red, green, and blue, after being normalized. After that, each feature was placed inside a 4 × 1 grid for each channel. This idea is illustrated in Figure 3, which shows a particular set of feature values.

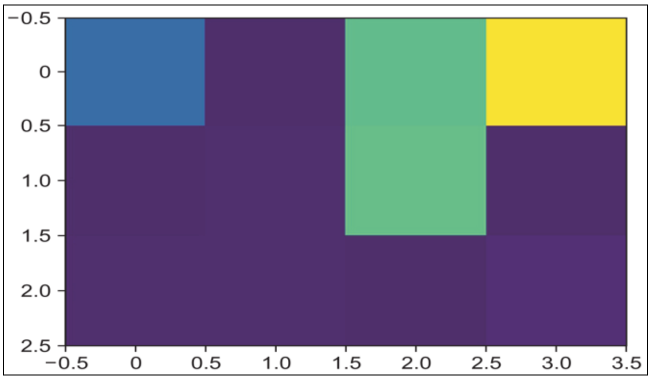

Fig 3: Image visualization for fatty liver disease dataset

While Figure 4(b) shows the minimum, maximum, and average validation accuracies during ten GA generations, Figure 4(a) shows the validation accuracy attained at each stage of the GA evolution. The highest hyper-parameter optimization accuracy was attained, according to an analysis of these numbers. On the other hand, a slightly different set of values for the filters and layers resulted in the lowest accuracy. These particular hyper-parameter setups show how they affect the validation accuracy of the model.

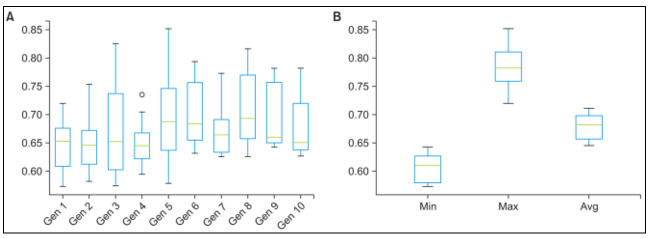

Fig 4: Validated accuracy (a) and summary of minimum, maximum, and average performance metrics of GA (b)

The performance of the hyper-parameter optimization utilizing the GA-based CNN during the training and testing stages is shown in Figure 5. Due to their demonstrated efficacy in earlier research, a number of machine learning methods, such as naïve Bayes, SVM, decision tree, logistic regression, and random forest, have been used for LD prediction.

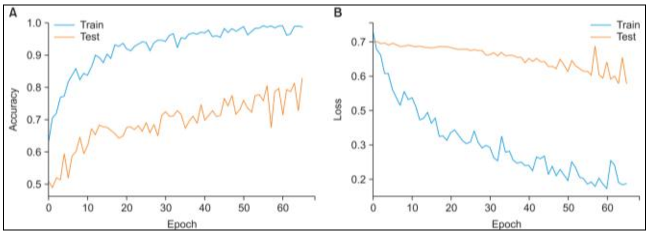

Fig 5: Performance outcomes of GA with ML on fatty liver disease (LD) prediction showing Loss and positive Accuracy

This study demonstrates that the prediction of LD is significantly improved by combining a GA with a CNN. With 100% accuracy, our method outperformed conventional feature engineering and demonstrated how well genetic algorithms can automate feature selection in complicated datasets. Our GA-based CNN model is significant because it has the potential to significantly improve public health by increasing the accuracy of LD detection. Better preventative measures and more individualized medical care may be made possible by this advancement. This method improves prediction accuracy and advances the creation of AI solutions for healthcare by using genetic algorithms for feature engineering. The model's success highlights how important interdisciplinary cooperation is to integrating these innovations into clinical practice, which eventually improves patient care and healthcare effectiveness.

5. Conclusion

Multi-institutional validation of GA-based CNN should be a top priority for future studies in order to verify its suitability and efficacy for a variety of patient populations. To improve and possibly raise the prediction accuracy, more research could look into synergies with other machine learning techniques. The model's sustainability and alignment with healthcare priorities will be determined in large part by ethical use of AI and cost-effectiveness assessments, which will advance AI's role in individualized patient care. This study concludes by highlighting how well GA-based feature engineering works to optimize CNNs for LD prediction.

6. References

1. Clark Jeanne M, Frederick Brancati L, Anna Mae Diehl. Nonalcoholic fatty liver disease. Gastroenterology. 2002; 122(6):1649-1657.
2. Muhammad Y, Tahir M, Hayat M, Chong KT. Early and accurate detection and diagnosis of heart disease using intelligent computational model. Sci Rep. 2020; 10(1):19747. Doi: https://doi.org/10.1038/s41598-020-76635-9
3. Zhang P, Xu F. Effect of AI deep learning techniques on possible complications and clinical nursing quality of patients with coronary heart disease. Food Sci Technol. 2022; 42:e42020. Doi: https://doi.org/10.1590/fst.42020
4. Elwahsh H, El-Shafeiy E, Alanazi S, Tawfeek MA. A new smart healthcare framework for real-time heart disease detection based on deep and machine learning. PeerJ Comput Sci. 2021; 7:e646. Doi: https://doi.org/10.7717/peerjcs.646
5. Al-Alshaikh HA, PP, Poonia RC, Saudagar AKJ, Yadav M, AlSagri HS, *et al*. Comprehensive evaluation and performance analysis of machine learning in heart disease prediction. Sci Rep. 2024; 14(1):7819. Doi: https://doi.org/10.1038/s41598-024-58489-7
6. Kanagarathinam K, Sankaran D, Manikandan R. Machine learning-based risk prediction model for cardiovascular disease using a hybrid dataset. Data Knowl Eng. 2022; 140:102042. Doi: https://doi.org/10.1016/j.datak.2022.102042
7. Theng D, Bhoyar KK. Feature selection techniques for machine learning: A survey of more than two decades of research. Knowl Inf Syst. 2024; 66:1575-1637. Doi: https://doi.org/10.1007/s10115-023-02010-5
8. Kurniabudi Stiawan D, Darmawijoyo, Bin Idris MY, Bamhdi AM, Budiarto R. CICIDS-2017 dataset feature analysis with information gain for anomaly detection. IEEE Access. 2020; 8:132911-132921. Doi: https://doi.org/10.1109/ACCESS.2020.3009843
9. Kaur S, Aggarwal H, Rani R. Hyper-parameter optimization of deep learning model for prediction of Parkinson's disease. Mach Vis Appl. 2020; 31:32. Doi: https://doi.org/10.1007/s00138-020-01078-1
10. Valarmathi R, Sheela T. Heart disease prediction using hyper parameter optimization (HPO) tuning. Biomed Signal Process Control. 2021; 70:103033. Doi: https://doi.org/10.1016/j.bspc.2021.103033
11. Montesinos-Lopez OA, Carter AH, Bernal-Sandoval DA, Cano-Paez B, Montesinos-Lopez A, Crossa J. A comparison between three tuning strategies for gaussian kernels in the context of univariate genomic prediction. Genes (Basel). 2022; 13(12):2282. Doi: https://doi.org/10.3390/genes13122282
12. Ghawi R, Pfeffer J. Efficient hyperparameter tuning with grid search for text categorization using kNN approach with BM25 similarity. Open Comput Sci. 2019; 9(1):160-180. Doi: https://doi.org/10.1515/comp-2019-0011
13. Kumar S, Ratnoo S. Multi-objective hyperparameter tuning of classifiers for disease diagnosis. Indian J Comput Sci Eng. 2021; 12(5):1334-1352. Doi: https://doi.org/10.21817/indjcse/2021/v12i5/211205081
14. Rusch T, Mair P, Hornik K. Structure-based hyperparameter selection with Bayesian optimization in multidimensional scaling. Stat Comput. 2023; 33(1):28. Doi: https://doi.org/10.1007/s11222-022-10197-w
15. Lankford S, Grimes D. Neural architecture search using particle swarm and ant colony optimization [Internet]. Ithaca (NY): arXiv.org; 2024 [cited at 2024 Jul 30]. Available from: http://arxiv.org/pdf/2403.03781
16. Tayebi M, El Kafhali S. Hyperparameter optimization using genetic algorithms to detect frauds transactions. Proceedings of the International Conference on Artificial Intelligence and Computer Vision (AICV2021); Cham, Switzerland: Springer, 2021, 288-297. Doi: https://doi.org/10.1007/978-3-030-76346-6_27
17. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, *et al*. Review of deep learning:

Concepts, CNN architectures, challenges, applications, future directions. J Big Data. 2021; 8(1):53. Doi: https://doi.org/10.1186/s40537-021-00444-8
18. Kong Y, Wang X, Cheng Y, Chen CL. Multi-stage convolutional broad learning with block diagonal constraint for hyperspectral image classification. Remote Sens. 2021; 13(17):3412. Doi: https://doi.org/10.3390/rs13173412
19. Ayan E. Genetic algorithm-based hyperparameter optimization for convolutional neural networks in the classification of crop pests. Arab J Sci Eng. 2024; 49:3079-3093. Doi: https://doi.org/10.1007/s13369-023-07916-4
20. Vetrimani E, Arulselvi M, Ramesh G. Building convolutional neural network parameters using genetic algorithm for the croup cough classification problem. Measur Sens. 2023; 27:100717. Doi: https://doi.org/10.1016/j.measen.2023.100717
21. Montecino DA, Perez CA, Bowyer KW. Two-level genetic algorithm for evolving convolutional neural networks for pattern recognition. IEEE Access. 2021; 9:126856-126872. Doi: https://doi.org/10.1109/ACCESS.2021.3111175
22. Salam A, Zeniarja J. Classification of deep learning convolutional neural network feature extraction for student graduation prediction. Indones J Electr Eng Comput Sci. 2023; 32(1):335. Doi: https://doi.org/10.11591/ijeecs.v32.i1.pp335-341
23. Zhu Y, Brettin T, Xia F, Partin A, Shukla M, Yoo H, *et al*. Converting tabular data into images for deep learning with convolutional neural networks. Sci Rep. 2021; 11(1):11325. Doi: https://doi.org/10.1038/s41598-021-90923-y
24. Hassan D, Hussein HI, Hassan MM. Heart disease prediction based on pre-trained deep neural networks combined with principal component analysis. Biomed Signal Proc Contr. 2023; 79(1):104019. Doi: https://doi.org/10.1016/j.bspc.2022.104019
25. Ahmad A, Feng C, Ge S, Yousif A. A survey on mining stack overflow: Question and answering (Q&A) community, Data Technologies and Applications. 2018); 52(2):190-247. Doi: https://doi.org/10.1108/dta-07-2017-0054, 2-s2.0-85051274733
26. Baltadzhieva A, Chrupała G. Question quality in community question answering forums, ACM SIGKDD Explorations Newsletter. 2015; 17(1):8-13. Doi: https://doi.org/10.1145/2830544.2830547
27. Binkhonain M, Zhao L. A review of machine learning algorithms for identification and classification of non-functional requirements, Expert Systems with Applications. 2019; 1. Doi: https://doi.org/10.1016/j.eswax.2019.100001, 2-s2.0-85070485369
28. Meth H, Brhel M, Maedche A. The state of the art in automated requirements elicitation, Information and Software Technology. 2013; 55(10):1695-1709. Doi: https://doi.org/10.1016/j.infsof.2013.03.008, 2-s2.0-84880850545
29. Iqbal T, Elahidoost P, Lúcio L. A bird's eye view on requirements engineering and machine learning, Proceedings of the 2018 25th Asia-Pacific Software Engineering Conference (APSEC), Nara, Japan, December 2018, 11-20. Doi: https://doi.org/10.1109/apsec.2018.00015, 2-s2.0-85066791675
30. Yousif A, Niu Z, Tarus JK, Ahmad A. A survey on sentiment analysis of scientific citations, Artificial Intelligence Review. 2019; 52(3):1805-1838. Doi: https://doi.org/10.1007/s10462-017-9597-8, 2-s2.0-85037615365
31. Ali H, Salleh MNM, Hussain K. *et al*. A review on data preprocessing methods for class imbalance problem, International Journal of Engineering & Technology. 2019; 8:390-397.
32. Wen J, Li S, Lin Z, Hu Y, Huang C. Systematic literature review of machine learning based software development effort estimation models, Information and Software Technology. 2012; 54(1):41-59. Doi: https://doi.org/10.1016/j.infsof.2011.09.002, 2-s2.0-80055086827
33. Magües DA, Castro JW, Acuna ST. HCI usability techniques in agile development, Proceedings of the 2016 IEEE International Conference on Automatica (ICA-ACCA), Curico, Chile, October 2016, 1-7. Doi: https://doi.org/10.1109/ica-acca.2016.7778513, 2-s2.0-85010433109
34. Achimugu P, Selamat A, Ibrahim R, Mahrin MNR. A systematic literature review of software requirements prioritization research, Information and Software Technology. 2014; 56(6):568-585. Doi: https://doi.org/10.1016/j.infsof.2014.02.001, 2-s2.0-84897033201
35. Lin Y-C, *et al*. Risk factors and predictors of non-alcoholic fatty liver disease in Taiwan, Ann. Hepatol, 2011.
36. Clark JM, *et al*., Defining nonalcoholic fatty liver disease: Implications for epidemiologic studies, Gastroenterology, 2003.
37. Savova GK, *et al*. Mayo clinic NLP system for patient smoking status identification, J. Am. Med. Inform. Assoc, 2008.
38. Le MH, *et al*. Prevalence of non-alcoholic fatty liver disease and risk factors for advanced fibrosis and mortality in the United States, PLoS One, 2017.