


RESEARCH

Open Access



Simulation of liver function enzymes as determinants of thyroidism: a novel ensemble machine learning approach

Abdullahi Garba Usman¹, Umar Muhammad Ghali², Mohamed Alhosen Ali Degm³, Salisu M. Muhammad⁴, Evren Hincal⁴, Abdulaziz Umar Kurya^{5*} , Selin Işık¹, Qendresa Hoti² and S. I. Abba^{6,7}

Abstract

Background: Hormone production by the thyroid gland is a prime aspect of maintaining body homeostasis. In this study, the ability of single artificial intelligence (AI)-based models, namely multi-layer perceptron (MLP), support vector machine (SVM), and Hammerstein–Weiner (HW) models, were used in the simulation of thyroidism status. The study's primary aim is to unveil the best performing model for the simulation of thyroidism status using hepatic enzymes and hormones as the independent variables. Three statistical metrics were used in evaluating the performance of the models, namely determination coefficient (R^2), correlation coefficient (R), and mean squared error (MSE).

Results: Considering the quantitative and visual presentation of the results obtained, it has been observed that the MLP model showed higher performance skills than SVM and HW, which improved their performances up to 3.77% and 12.54%, respectively, in the testing stages. Furthermore, to boost the performance of the single AI-based models, three different ensemble approaches were employed, including neural network ensemble (NNE), weighted average ensemble (WAE), and simple average ensemble (SAE). The quantitative predictive performance of the NNE technique boosts the performance of SAE and WAE approaches up to 2.85% and 1.22%, respectively, in the testing stage.

Conclusions: Comparative performance of the ensemble techniques over the single models showed that NNE outperformed all the three AI-based models (MLP, SVM, and HW) and boosted their performance accuracy up to 7.44%, 11.212%, and 19.98%, respectively, in the testing stages.

Keyword: Thyroidism, AI-based models, Ensemble machine learning, Liver function enzymes, Hormones

Background

Hormone production by the thyroid gland is a prime aspect of maintaining body homeostasis. However, the process must be tightly regulated to prevent a hormonal imbalance that negatively affects certain metabolic activities associated with many diseases (Ghali et al. 2020). The thyroid gland is part of the endocrine system, which produces thyroid hormones, namely triiodothyronine (T_3)

and thyroxine (T_4), responsible for hormonal regulation for proper functioning (Brent 2012).

The thyroid-stimulating hormone produced by the pituitary gland is essential in regulating hormonal production by the thyroid gland (Jonklaas 2020). Clinically, an elevated level of TSH denotes underperformance of thyroid hormones; as such, the pituitary gland compensates accordingly by producing more TSH, a condition referred to as hypothyroidism. However, low TSH levels indicate excess production of thyroid hormone above normal. The pituitary hormone compensates accordingly by decreasing the TSH production to retain the thyroid

*Correspondence: kurya360@gmail.com

⁵ School of Life and Allied Health Sciences, Department of Biotechnology, Glocal University, Saharanpur, Uttar Pradesh 247121, India
Full list of author information is available at the end of the article

function, a condition referred to as hyperthyroidism (Zhang et al. 2020).

Considering that our research is aimed at simulation of thyroidism status using liver enzymes and thyroid-stimulating hormone as an independent variable, we feel it is worthy to further elucidate the biological relevance and association of thyroid hormones and vitamin D with disease severity and liver functions in patients with non-cholesteric chronic liver disease (Fisher and Fisher 2007). More so, thyroid hormones play a substantial role in the biological system. However, their abnormal level leads to several medical implications such as thyroid tumorigenesis, Hashimoto's thyroiditis, goitre, myxoedema, benign tumour of the pituitary gland, peripheral neuropathy, as observed in many patients (Choi et al. 2021; Huang and Liaw 1995). Abnormalities in serum liver enzymes are observed in hypothyroidism and may be related to impaired lipid metabolism, hepatic steatosis, hypothyroidism-induced myopathy, hyperammonemia, ascites, and hepatitis c virus. (Piantanida et al. 2020; Płudowski et al. 2013). We believe that proposing the best performing AI and machine learning model using hepatic enzymes and hormones will serve as a breakthrough in drug design and discovery for several medical complications.

Interestingly, machine learning tools were widely employed to extract patterns within patient data and predict the outcomes for improved clinical management of autoimmune thyroid diseases. Recently, a comparison of logistic regression and neural network models was conducted to diagnose thyroid disorders with better performance of the neural network model than multinomial logistic regression. However, the predictive performance of the two models was found to be disdainful to laboratory variables (Chowdhury and Chakraborty 2017).

Another study that uses data set from the Graven Institute in Sydney, Australia, to predict thyroid disease using a machine learning approach shows that prediction and classification of any data depend on the data set. The algorithms used, such as SVM, Naïve Bayes, autoencoders, ANNs, and CNNs yield a better result (Chaubey et al. 2020). A cross-sectional study of 143 LN patients with hyperthyroidism diagnosed by renal biopsy using PCA–logistic regression model shows that the model performed well in identifying important risk factors for specific clinical outcomes (Aguayo-Orozco et al. 2021).

Association between hyperthyroidism and liver disease has been reported and further correlates hyperthyroidism's severity as an independent risk factor for abnormal liver function enzymes (Piantanida et al. 2020). Furthermore, many subsequent case reports and series have highlighted the prevalence of liver test abnormalities in the setting of hyperthyroidism and several mechanisms

of liver dysfunction in hyperthyroidism, including liver abnormalities due to hyperthyroidism alone, liver damage related to heart failure and hyperthyroidism, and concomitant liver disease in hyperthyroidism (Punekar et al. 2018). Moreso, an experimental study involving the application of linear and nonlinear models has predicted the thyroid hormone level (TSH) using different macro-elements and vitamins as the corresponding input parameters (Muhammad Ghali et al. 2020). In a clinical puzzle study, liver transplant evaluation was carefully analysed in a patient with coagulopathy, encephalopathy, and drug-induced acute liver failure. Uncontrolled thyroid diseases are reported to strongly correlate with liver dysfunction (Anugwom and Leventhal 2021).

Unfortunately, none of the studies addresses the correlations and associations with artificial intelligence and machine learning using two independent variables. The major novelty of our work is the adequate use of machine learning algorithms to understand such complex clinical data better and propose the best AI algorithm that could model thyroidism status in different subjects using liver function enzymes and level of TSH as an input variable. Interestingly, this compelling evidence in our work suggested that the best performed model has tremendous potential in the future for further endocrine-based research, diagnosis, and treatment of several liver diseases.

The principal aim of our work is to make a comparative analysis and propose the best performing model of AI for the simulation of thyroidism status using hepatic enzymes and hormones as the independent variables, considering the gaps and limitations in many published articles.

Methods

Clinical methodology

The sample was assayed as described by (Muhammad Ghali et al. 2020). Briefly, an automated COBAS E411 analyser was applied to analyse liver function enzymes. Thyroid hormones were assessed using Elecsys COBAS E411 after centrifuging it for 10 min at 2000g in a red-capped tube. The samples were separated into Control groups, hyperthyroidism and hypothyroidism.

Proposed data computational intelligence approach

Different data-driven approaches were used separately in this research to propose a model for developing an AI diagnostic approach for TSH. This current study is data-driven, collecting the data from our previous research (Ghali et al. 2020). The thyroidism status was predicted using two parameters that serve as inputs, i.e. liver enzymes alanine transaminase (ALT), aspartate transaminase (AST), albumin blood test (ALB), gamma-glutamyl

transferase (GGT), alkaline phosphatase (ALP), direct bilirubin (DBIL), total bilirubin (TBIL), and hormones such as thyroid-stimulating hormone (TSH), triiodothyronine (T3), thyroxine (T4), free triiodothyronine (FT3), and free thyroxine (FT4). The liver enzyme parameters were used to predict thyroidism status using different AI models, considering that liver is the major organ that synthesizes thyroid-binding globulin, prealbumin, and albumin that binds to thyroid hormone in the peripheral circulation and the liver metabolizes thyroid hormone. Therefore, to obtain accurate predictions of the best performing AI model, liver function enzyme parameters are likely to give a clear picture of the thyroidism status since the liver plays a crucial role in thyroid disease conditions.

Furthermore, three ensemble techniques (MLP, SVM, and HW) were applied to boost the prediction accuracy of the thyroidism status by combining the output results of the single models. Practically, concluding on a single model that can outperform other existing models used in predicting various parameters for a specific study is not feasible for the predictors. The proposed method in this research involves the determination of thyroid hormone status using two liver enzymes and hormones by selecting an ensemble of different models.

Hammerstein–Weiner model (HW)

The Hammerstein–Weiner (HW) model involves utilizing a black-box discovery model methodology planned to decide the nonlinear framework (Gaya et al. 2017). The HW model's arrangement consists of three blocks: a static input nonlinear block, a static output nonlinear, and the linear dynamic block, as shown in Fig. 1 (Abba et al. 2019). The model adopts the nonlinear input to linear function blocks, then returns to nonlinear functions in the output structure. Furthermore, the HW model displays more precise comprehension of the nonlinear and linear system connection than the other standard ANNs (Abba et al. 2019). Mainly, the MATLAB toolbox was utilized to improve the HW model based on its structure. The piecewise linear functions are input and the output

nonlinearity predictors, while 10 is set as default for the number of units, although the complexity of the model increases as the number of units gets more extensive (Guo 2004).

Multi-layer perceptron (MLP) neural network

Multi-layer perceptron (MLP) is considered among the frequent exemplars of ANN, which carry a nonlinear system and work as a widespread approximator (Choubin et al. 2016). The structure of the MLP consists of output, input, and hidden layer, unlike the other ordinary ANNs (Kim and Singh 2014; Pham et al. 2019). The input layer nodes are mainly linked to the hidden and output layers. From the input to the output layer, the signals are processed and afterwards transmitted across the assist of biases and weights by sequential mathematical operations. The Levenberg–Marquardt algorithm is a learning algorithm that is mainly used to improve the inaccuracy among the measured and predicted values. The training algorithms are constantly replicated till the required outcomes are pointed out. The MLP structure consists of input, output, and one or more hidden layers, similar to the ordinary ANNs, shown in Fig. 2 (Committee 2000).

$$y_i = \sum_{j=1}^N w_{ji}x_j + w_{i0} \quad (1)$$

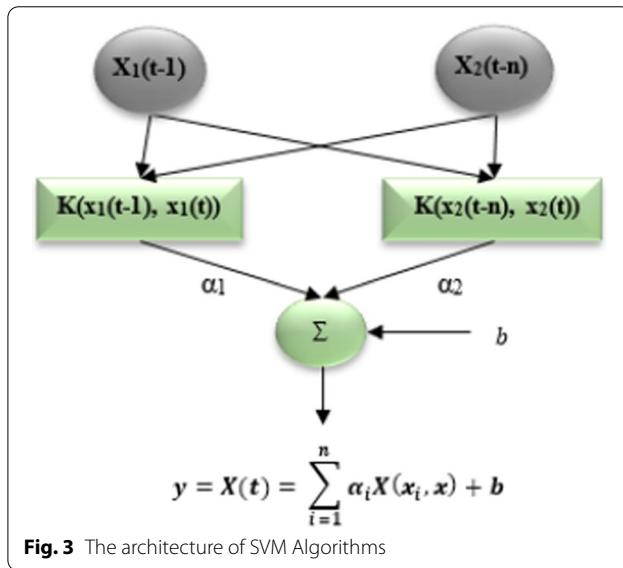
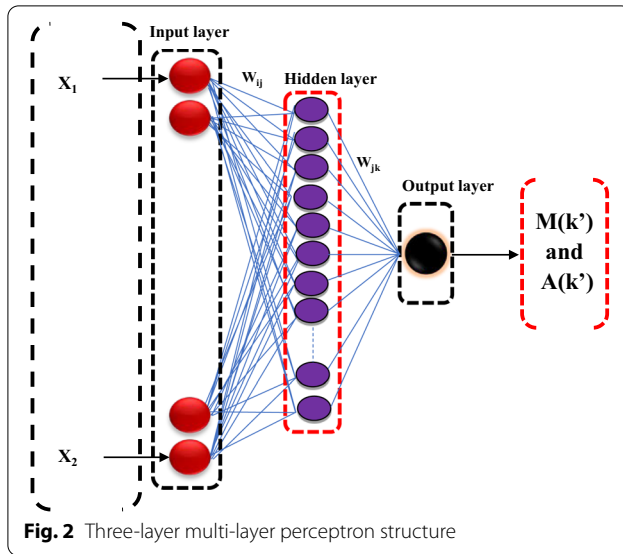
N is defined as the total number of nodes in the top layer of the node, i ; w_{ji} is the weight between the nodes i and j in the upper layer; x_j defines the output derived from node j ; w_{i0} is the bias in node i , and y_i defines the input signal of node i which crosses via the transfer function.

Support vector machine (SVM)

The concept of learning in the theme of support vector machine (SVM) was suggested by Vapnik in 1995, which supplies the wanted machine for problem-solving that include prediction, classification, regression, and pattern recognition. The SVM works and consists of a data-driven model. The two prominent roles of SVM include



Fig. 1 Schematic of Hammerstein–Wiener model. $w(t) = f(u(t))$ is the nonlinear function that converts input data, $x(t) = w(t)B/F$ is a linear transfer function, f and h are acting on the input and output port of the linear block, respectively. The other functions, $w(t)$ and $x(t)$ are used to define the output and input of the linear block



statistical learning theory and structural risk minimization. SVM's ability to boost the overall efficacy of the model and decrease the error, excess of data, and sophistication as well the development in the overall efficacy of the system, makes it superior to ANN (Vapnik 1995).

SVM can be categorized into linear support vector regression and nonlinear support vector regression. This indicates that support vector regression (SVR) is estimated as a category of SVM according to the two primary structural layers; the kernel function weighting on the input variable as the first layer, and the second function is a kernel outputs weighted sum as demonstrated in Fig. 3. The linear regression adapts on the data; afterwards, the

outputs pass over the nonlinear kernel to capture the nonlinear model of the data.

Ensemble techniques

The AI-based models for the same inputs provide diverse performance levels regarding their robustness or limitation. Hence, many methods such as web ranking algorithm, classification, regression problems, and time series clustering are used in different study fields (B and Sad-aoui 2019; Baba et al. 2015a; Dehghanian et al. 2015; Loos et al. 2019). Ensemble learning is the collective term for branch machine learning that deal with homogenous or heterogenous multiple models. The ensemble machine learning method is usually engaged by joining the process of various predictors to boost the performance of a single AI model. Machine learning has been demonstrated to be exceptionally successful in creating exact outcomes compared to single models applied for tackling a similar issue. For developing the expected performance of this model, three procedures are commonly utilized: (1) simple averaging ensemble (SAE) for combining the HW, MLP, and SVM predictors, (2) neural network ensemble (NNE), and (3) weighted averaging ensemble (WAE) (Baba et al. 2015b).

Simple averaging ensemble (SAE)

For SAE, the SVM, HW, and MLP single models are first prepared and tried independently, the average of the MLP, SVM, and HW are analysed and tested against the noticed qualities where the overall formula for SAE is given as:

$$P(t) = \frac{1}{N} \sum_{i=1}^N p_i(t) \quad (2)$$

N defines the number of learners (here $N=3$), and p_i represents the output of any single model (i.e. HW, SVM, and MLP) at a specific time t .

Weighted average ensemble (WAE)

Weighted average ensemble (WAE) can be resolved by allocating different weights to specific outputs of the single models indicated to their significance output, which is the opposite in the case of single models. The WAE is shown in the form of:

$$P(t) = \sum_{i=1}^N w_i p_i(t) \quad (3)$$

where w_i represents the weight applied to the i th model's output and can be resolved based on the performance of the model as:

$$w_i = \frac{DC_i}{\sum_{i=1}^N DC_i} \quad (4)$$

DC_i is the performance efficiency of the i th single model.

Neural network ensemble (NNE)

In the neural ensemble method (NNE), the nonlinear average is directed by training the different neural networks. The input layer of NNE is supported through outputs of the single models, by which everyone in the input layer is assigned to a single neuron. For network training, the backpropagation algorithm is used, by which the perfect structure and epoch number can be demonstrated utilizing the trial by error method for the ensemble network.

Data pre-processing and model validation

In computational intelligence models, the principal point is to guarantee any particular model or models utilized upon a given data collection and achieve agreeable predictions on obscure data collection (Nourani et al. 2018). The most frequent issue in prediction is overfitting, which results in the contradiction between testing performances and training. Different validation methods can be applied during the validation process, such as k-fold cross-validation and leave one out and holdout. The primary importance of k-fold and cross-validation is that at every round, the validation and training set are autonomous from each other (Usman et al. 2020). In our study, the k-fold cross-validation is used to adapt and reduce overfitting issues as demonstrated in Fig. 4.

Furthermore, the primer training data set is separated into same-sized subsets of k and typed from the $k-1$ data subsets for the validation process, while the remaining subsets are used for the training purpose (Elkiran et al. 2018). The result variations are considered as the average of validation efficiency of k -subsets. In general, k -values are calculated from sample availability, mainly

2–10. In the k-fold cross-validation process, the general advantages are that the calibration and validation set in every round are independent of one other to achieve a satisfying foundation of model optimization (Abba et al. 2017). The basic set of data is split into two groups; the verification and calibration set to achieve high performance of the data usage in model configuration (Soltani et al. 2015). Our study conducted data classification in two phases (25% for verification and 75% for calibration) to avoid the overfitting, underfitting, and local minima issues that may lead to qualitative and quantitative changes, as shown in Fig. 4 (Usman et al. 2021).

The performance accuracy is estimated from various criteria based on the differential between predicted and measured values. In our study, correlation coefficient (R), determination coefficient (R^2), and mean square error (MSE) were used to evaluate the models:

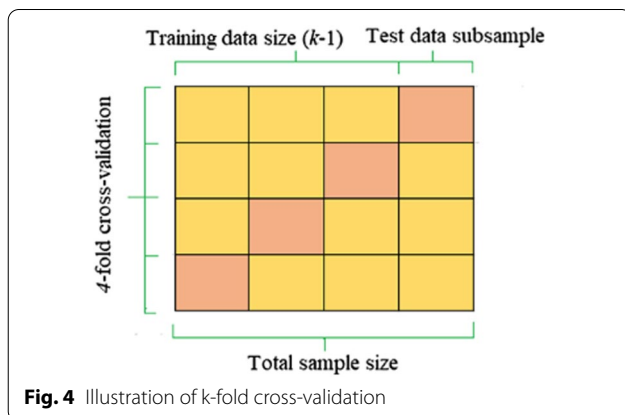
$$R^2 = 1 - \frac{\sum_{j=1}^N [(Y_{obsi} - Y_{comi})]^2}{\sum_{j=1}^N [(Y_{obsi} - \bar{Y}_{obsi})]^2} \quad (5)$$

$$MSE = \frac{1}{N} \sum_{(i=1)}^N (Y_{obsi} - Y_{comi})^2 \quad (6)$$

$$R = \frac{\sum_{i=1}^N (Y_{obsi} - \bar{Y}_{obsi})(Y_{comi} - \bar{Y}_{comi})}{\sqrt{\sum_{i=1}^N (Y_{obsi} - \bar{Y}_{obsi})^2 \sum_{i=1}^N (Y_{comi} - \bar{Y}_{comi})^2}} \quad (7)$$

where N =data number, Y_{obsi} =observed data, \bar{Y} = average value, and Y_{comi} = computed values.

According to Nourani et al. (2018) and Elkiran et al. (2019), for a good analysis of any data intelligence model, the efficiency performance should include at least one goodness-of-fit (e.g. R^2) and at least one absolute error measure (e.g. RMSE). The employed three performance criteria in this study were attributed to the fact that multi-criteria indicator for measuring the models' performance was generally employed in contemporary studies. Another important reason for using multiple criteria is that the properties of data, such as normality, size, and linearity, affect the performance accuracy of any model, which can also be evaluated using these criteria. In addition, several studies have already shown that even for the same type of data set, the performance results may deviate from one model performance to another. For example, R^2 does not consider any biases that might be present in the data. Therefore, a good model might have a low R^2 value or a model that does not fit the data might have a high R^2 value. Hence, other evaluation metrics can be combined with the goodness-of-fit (R^2), such as the error measure root mean square (RMSE) and biases measure



could lead to promising and reliable simulation. Other performance efficiency criteria can also be used, such as mean absolute relative error (MAE) (Nourani et al. 2018; Elkiran et al. 2018; Usman et al. 2021).

Results

The efficiency of the AI-based models was checked by comparing the predicted and experimental values regarding the thyroid status of the subjects. The performance skills of the techniques used in the determination were estimated using three different statistical metrics (R^2 , MSE, and R). Furthermore, visual presentation of the data via scatter plots, radar plots, bar charts, and response plots was equally illustrated. The obtained results are shown in the following section.

Performance of the single AI-based models and ensemble techniques in both the training and testing stages in terms of the statistical evaluation metrics are shown in Tables 1 and 2. The tables equally demonstrate the quantitative comparative performance of the models towards simulating the dependent variable inform of thyroidism status using various liver function enzymes and hormones as the independent variables.

Table 2 Performance of the single models

| Models | R^2 | R | MSE |
|-----------------|--------|--------|--------|
| <i>Training</i> | | | |
| MLP | 0.9331 | 0.9660 | 0.0162 |
| SVM | 0.8939 | 0.9455 | 0.0258 |
| HW | 0.8085 | 0.8992 | 0.0465 |
| <i>Testing</i> | | | |
| MLP | 0.9254 | 0.9620 | 0.0198 |
| SVM | 0.8877 | 0.9422 | 0.0290 |
| HW | 0.8000 | 0.8944 | 0.0500 |

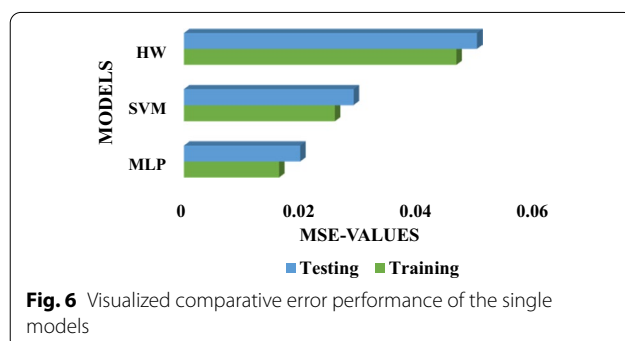
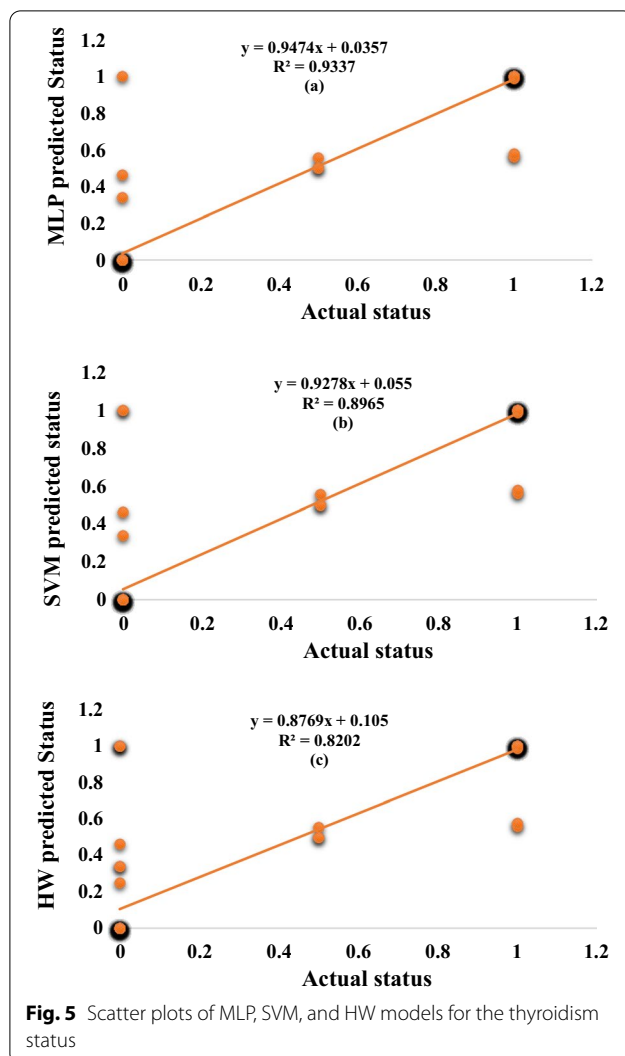
The result for the comparative performance of the models is equally shown in Fig. 5 using a scatter plot to recommend the best performing algorithm in both the training and testing stages.

The errors depicted by each model is demonstrated comparatively in Fig. 6 to visualize the performance results of the models; this is in line with the quantitative performance result shown in Table 1.

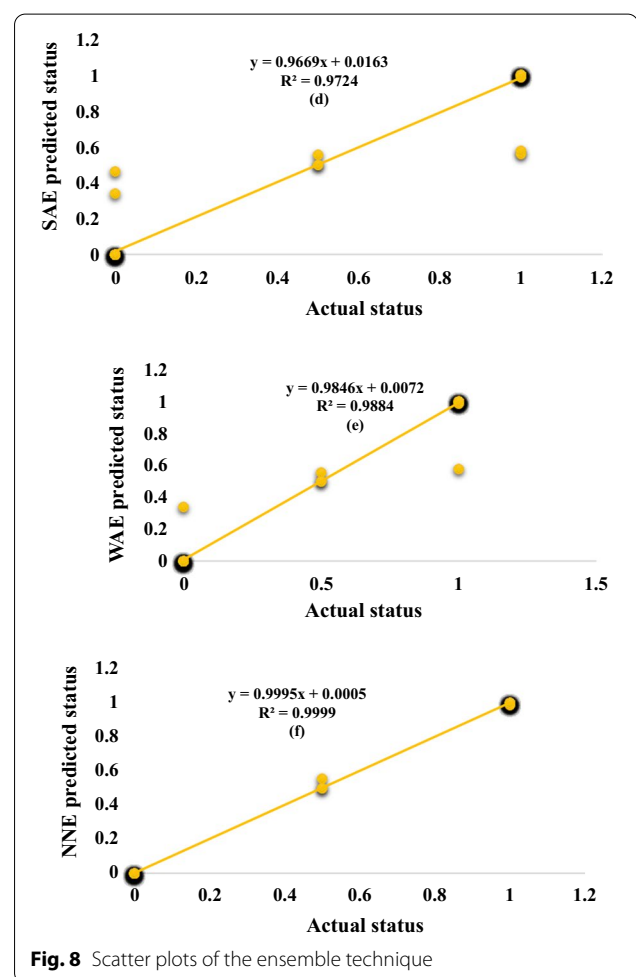
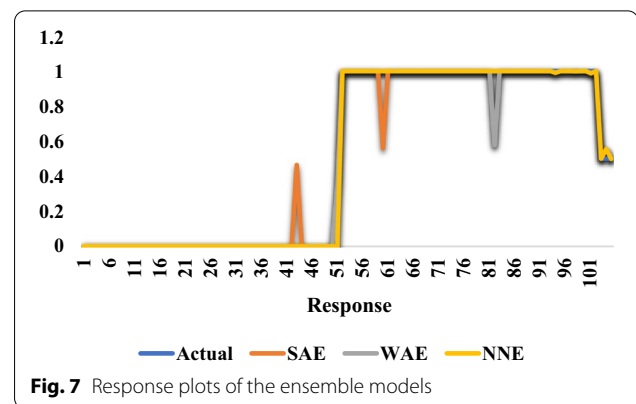
Furthermore, the NNE technique boosted the performance accuracy of the single AI models MLP, SVM,

Table 1 Basic descriptive statistics and Spearman–Pearson correlation analysis

| Variables | ALB | D.BIL | T.BIL | GGT | ALP | AST | ALT | FT4 | FT3 | T4 | T3 | TSH | Status |
|--------------------|-------|---------|-------|-------|-------|-------|-------|-------|-------|---------|-------|-------|--------|
| Mean | 3.6 | 0.2 | 0.6 | 27.9 | 129.1 | 32.5 | 31.8 | 20.9 | 5.2 | 125.9 | 3.6 | 9.0 | 2.0 |
| Median | 3.5 | 0.2 | 0.5 | 24.0 | 118.5 | 32.0 | 31.0 | 15.4 | 4.2 | 113.0 | 2.0 | 1.8 | 2.0 |
| Mode | 3.5 | 0.1 | 0.5 | 21.0 | 89.0 | 34.0 | 39.0 | 16.6 | 1.9 | 113.0 | 1.7 | 0.0 | 1.0 |
| Standard Deviation | 0.7 | 0.1 | 0.2 | 11.7 | 52.8 | 9.6 | 12.1 | 16.9 | 3.9 | 86.2 | 4.7 | 20.3 | 0.8 |
| Kurtosis | 1.0 | 9.5 | 0.5 | 0.0 | 0.6 | − 0.1 | − 0.2 | 2.4 | 0.2 | − 0.9 | 31.1 | 13.4 | − 1.5 |
| Skewness | 0.6 | 2.4 | 0.7 | 0.9 | 0.9 | 0.3 | 0.4 | 1.6 | 0.9 | 0.4 | 4.9 | 3.6 | 0.0 |
| Range | 4.0 | 0.6 | 1.2 | 55.4 | 259.0 | 52.0 | 59.3 | 83.3 | 16.3 | 318.2 | 36.7 | 108.0 | 2.0 |
| Minimum | 2.1 | 0.1 | 0.1 | 0.6 | 36.0 | 11.0 | 9.7 | 2.0 | 0.1 | 0.4 | 0.1 | 0.0 | 1.0 |
| Maximum | 6.1 | 0.7 | 1.3 | 56.0 | 295.0 | 63.0 | 69.0 | 85.3 | 16.4 | 318.6 | 36.9 | 108.0 | 3.0 |
| Variables | ALB | D.BIL | T.BIL | GGT | ALP | AST | ALT | FT4 | FT3 | T4 | T3 | TSH | Status |
| ALB | 1 | | | | | | | | | | | | |
| D.BIL | 0.160 | 1.000 | | | | | | | | | | | |
| T.BIL | 0.040 | 0.196 | 1.000 | | | | | | | | | | |
| GGT | 0.148 | 0.209 | 0.167 | 1.000 | | | | | | | | | |
| ALP | 0.066 | 0.129 | 0.004 | 0.345 | 1.000 | | | | | | | | |
| AST | 0.003 | − 0.021 | 0.025 | 0.275 | 0.205 | 1.000 | | | | | | | |
| ALT | 0.025 | − 0.049 | 0.081 | 0.207 | 0.164 | 0.663 | 1.000 | | | | | | |
| FT4 | 0.104 | − 0.003 | 0.080 | 0.252 | 0.069 | 0.189 | 0.304 | 1.000 | | | | | |
| FT3 | 0.075 | 0.038 | 0.119 | 0.274 | 0.060 | 0.273 | 0.344 | 0.816 | 1.000 | | | | |
| T4 | 0.059 | − 0.024 | 0.066 | 0.265 | 0.089 | 0.255 | 0.399 | 0.847 | 0.831 | 1.000 | | | |
| T3 | 0.003 | 0.025 | 0.174 | 0.163 | 0.028 | 0.049 | 0.122 | 0.667 | 0.632 | 0.644 | 1.000 | | |
| TSH | 0.092 | 0.047 | 0.154 | 0.005 | 0.001 | 0.029 | 0.170 | 0.375 | 0.441 | − 0.494 | 0.240 | 1.000 | |
| Status | 0.003 | − 0.028 | 0.093 | 0.277 | 0.039 | 0.220 | 0.312 | 0.765 | 0.820 | 0.898 | 0.569 | 0.503 | 1 |

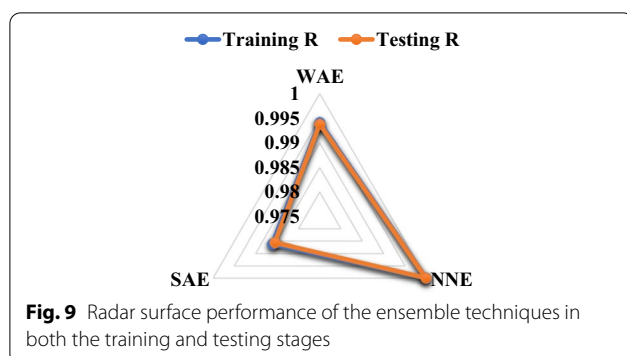


and HW up to 7.44%, 11.21%, and 19.98%, respectively, in the testing stages. Besides, the comparative performance of the models is illustrated graphically, as shown in Figs. 7, 8, and 9.



Discussion

Spearman correlation analysis shown in Table 3 depicts a strong correlation between the dependent variable inform of thyroidism status with TSH, T3, FT3, and T4 having correlation coefficient values as 0.503, 0.569, 0.765, 0.820, and 0.898, respectively. It equally showed

**Table 3** Performance of the ensemble techniques

| Models | R^2 | R | MSE |
|-----------------|--------|--------|--------|
| <i>Training</i> | | | |
| SAE | 0.9724 | 0.9861 | 0.0067 |
| WAE | 0.9884 | 0.9942 | 0.0028 |
| NNE | 0.9999 | 0.9999 | 0.0000 |
| <i>Testing</i> | | | |
| SAE | 0.9713 | 0.9855 | 0.0689 |
| WAE | 0.9876 | 0.9938 | 0.0032 |
| NNE | 0.9998 | 0.9999 | 0.0000 |

a weak inverse correlation between the dependent variable and D.BIL. The basic descriptive statistics equally demonstrate the nature of the raw data before the data normalization and the modelling, which will aid in simplifying the approach to be employed during the simulation. More information regarding correlation analysis can be found in (Abba et al. 2021; Asnake Metekia et al. 2021).

Based on the performance of the models, it can be observed that MLP outperformed both SVM and HW models in both the training and testing stages. Furthermore, the quantitative skills of MLP showed it could improve the performance ability of HW SVM and HW up to 3.77% and 12.54, respectively, in the testing stages.

Based on the quantitative performance of the ensemble techniques, it can be noted that all three approaches were able to predict the dependent variable with high performance in both the training and testing stages. Moreover, the NNE method outperformed the other two techniques (SAE and WAE) in the training and testing phases. Moreover, the quantitative predictive performance of the NNE technique can boost the performance of SAE and WAE approaches up to 2.85% and 1.22%, respectively, in the testing stage.

The comparative performance of the ensemble techniques can be presented visually using a radar plot using

the correlation coefficient of the ensemble techniques in both the training and testing stages. Based on this plot, it can be noted that the NNE approach outperformed SAE and WAE techniques in both the training and testing stages, as shown in Table 2.

Furthermore, other approaches such as the hybrid data intelligence techniques that involve the coupling of both the linear and nonlinear models to enjoy the benefits of both properties towards enhancing the performance skills of the single models as well as the implementation of the emerging and newest metaheuristic approaches such as the Harris Hawks optimization method and the novel neuro-emotional technique can be used to improve and boost the performance efficiency of the single models.

Conclusions

The results obtained from the data-driven approaches using experimental data indicated that MLP, SVM, and HW AI models could model the thyroidism status of each subject either as normal, hypothyroidism, or hyperthyroidism with R^2 -values higher than 0.7000 in both the training and testing stages. Interestingly, NNE model outperformed all the three AI-based models (MLP, SVM, and HW) and boosted their performance accuracy up to 7.44%, 11.212%, and 19.98%, respectively, in the testing stages, which makes it the most recommended and promising model for further endocrine-based research. Therefore, our study serves as a breakthrough and calls for the application of the proposed model in drug design and discovery to reduce the morbidity and mortality rate of hepatocytes and thyroid diseases.

Abbreviations

AI: Artificial intelligence; ALT: Alanine transaminase; AST: Aspartate transaminase; ALB: Albumin blood test; ALP: Alkaline phosphatase; DBIL: Direct bilirubin; ECF: Extracellular fluid; FT₃: Free triiodothyronine; FT₄: Free thyroxine; GGT: Gamma-glutamyl transferase; HW: Hammerstein–Weiner; MSE: Mean squared error; MLP: Multi-layer perceptron; NNE: Neural network ensemble; PTH: Parathyroid hormone; R: Correlation coefficient; R^2 : Determination coefficient; SAE: Simple average ensemble; SVM: Support vector machine; TSH: Thyroid-stimulating hormone; TBIL: Total bilirubin; T₃: Triiodothyronine; TSH: Thyroid-stimulating hormone; T₄: Thyroxine; WAE: Weighted average ensemble.

Acknowledgements

The authors are grateful to Near East University and the relevant cited articles used in this manuscript.

Authors' contributions

AGU and UMG contributed to the conceptualization, analysis and writing of the manuscript. MAA, SMM and EH contributed with the proofreading. AUK and QH contributed to writing the manuscript. SI helps towards proofreading and major revisions of the manuscript, while SIA supervised and drafted the manuscript. All the authors participated in the proofreading and approval of the manuscript.

Funding

Not applicable.

Availability of data and materials

The data are provided in the main manuscript.

Declarations**Ethics approval and consent to participate**

Not applicable.

Consent for publication

Not applicable.

Competing interests

There is no known competing interest to be declared by the authors.

Author details

¹Faculty of Pharmacy, Department of Analytical Chemistry, Near East University, 99138 Nicosia, Turkish Republic of Northern Cyprus. ²Faculty of Medicine, Department of Medical Biochemistry, Near East University, 99138 Nicosia, Turkish Republic of Northern Cyprus. ³National Centre for Diabetes and Endocrinology, Tripoli, Libya. ⁴Department of Mathematics, Near East University, 99138 Nicosia, Turkish Republic of Northern Cyprus. ⁵School of Life and Allied Health Sciences, Department of Biotechnology, Global University, Saharanpur, Uttar Pradesh 247121, India. ⁶Faculty of Engineering, Department of Civil Engineering, Baze University, Abuja, Nigeria. ⁷Interdisciplinary Research Center for Membrane and Water Security, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia.

Received: 17 January 2022 Accepted: 4 March 2022

Published online: 21 March 2022

References

- Abba SI, Abdulkadir RA, Sammen SSH, Usman AG, Meshram SG, Malik A, Shahid S (2021) Comparative implementation between neuro-emotional genetic algorithm and novel ensemble computing techniques for modelling dissolved oxygen concentration. *Hydrol Sci J*. <https://doi.org/10.1080/02626667.2021.1937179>
- Abba SI, Hadi SJ, Abdullahi J (2017) River water modelling prediction using multi-linear regression, artificial neural network, and adaptive neuro-fuzzy inference system techniques. *Procedia Comput Sci* 120:75–82. <https://doi.org/10.1016/j.procs.2017.11.212>
- Abba SI, Nourani V, Elkiran G (2019) Multi-parametric modeling of water treatment plant using AI-based non-linear ensemble. 2, 1–15. <https://doi.org/10.2166/wst.2011.079>
- Abba SI, Saleh A, Hamza N, Tukur AI, Wahab NA (2019a) Modelling of uncertain system: a comparison study of linear and non-linear approaches. *IEEE*
- Abba SI, Saleh A, Hamza N, Tukur AI, Wahab NA (2019b) Modelling of uncertain system: a comparison study of linear and non-linear approaches. *IEEE*
- Aguayo-Orozco A, Brunak S, Taboureaux O (2021) Extrapolation of drug induced liver injury responses from cancer cell lines using machine learning approaches. *Comput Toxicol*. <https://doi.org/10.1016/j.comtox.2020.100147>
- Anugwom CM, Leventhal TM (2021) Thyroid disease-induced hepatic dysfunction: a clinical puzzle. *ACG Case Rep J* 8(4):e00555. <https://doi.org/10.14309/crj.0000000000000555>
- Asnake Metekia W, Garba Usman A, Hatice Ulusoy B, Isah Abba S, Chirkena Bali K (2021) Artificial intelligence-based approaches for modeling the effects of spirulina growth mediums on total phenolic compounds. *Saudi J Biol Sci*. <https://doi.org/10.1016/j.sjbs.2021.09.055>
- B FA, Sadaoui S (2019) Multi-class ensemble learning. Springer International Publishing. <https://doi.org/10.1007/978-3-030-18305-9>
- Baba NM, Makhtar M, Abdullah S, Awang MK (2015) Current Issues in Ensemble Methods and Its Applications 81(2):266–276
- Brent GA (2012) Science in Medicine Mechanisms of Thyroid Hormone Action 122(9):3035–3043. <https://doi.org/10.1172/JCI60047.three>
- Chaube G, Bisen D, Arjaria S, Yadav V (2020) Thyroid disease prediction using machine learning approaches. *Natl Acad Sci Lett*. <https://doi.org/10.1007/s40009-020-00979-z>
- Choi SY, Yi DY, Kim SC, Kang B, Choe BH, Lee Y, Lee YM, Lee EH, Jang HJ, Choi YJ, Kim HJ (2021) Severe phenotype of non-alcoholic fatty liver disease in pediatric patients with subclinical hypothyroidism: a retrospective multicenter study from Korea. *J Korean Med Sci* 36(20):1–10. <https://doi.org/10.3346/jkms.2021.36.e137>
- Choubin B, Khalighi-Sigaroodi S, Malekian A, Kişi Ö (2016) Multiple linear regression, multi-layer perceptron network and adaptive neuro-fuzzy inference system for forecasting precipitation based on large-scale climate signals. *Hydrol Sci J* 61(6):1001–1009. <https://doi.org/10.1080/02626667.2014.966721>
- Chowdhury S, Chakraborty P, pratim. (2017) Universal health coverage - There is more to it than meets the eye. *J Family Med Prim Care* 6(2):169–170. <https://doi.org/10.4103/jfmpc.jfmpc>
- Committee AT (2000) Artificial neural networks in hydrology. I: Preliminary concepts. *J Hydrol Eng* 5(2):115–123
- Dehghanian E, Kaykhaii M, Mehrpur M (2015) Comparison of single best artificial neural network and neural network ensemble in modeling of palladium microextraction. *Monatshefte Fur Chemie* 146(8):1217–1227. <https://doi.org/10.1007/s00706-014-1396-1>
- Elkiran G, Nourani V, Abba SI, Abdullahi J (2018) Artificial intelligence-based approaches for multi-station modelling of dissolve oxygen in river. *Glob J Environ Sci Manag* 4(4):439–450. <https://doi.org/10.22034/gjesm.2018.04.005>
- Fisher L, Fisher A (2007) Vitamin D and parathyroid hormone in outpatients with noncholestatic chronic liver disease. *Clin Gastroenterol Hepatol* 5(4):513–520. <https://doi.org/10.1016/j.cgh.2006.10.015>
- Gaya MS, Yusuf LA, Mustapha M, Muhammad B, Sani A, Tijjani A, Khairi MTM (2017) Estimation of turbidity in water treatment plant using Hammerstein–Wiener and neural network technique. *Indonesian J Electr Eng Comput Sci* 5(3):666–672
- Ghali UM, Alhosen M, Degm A, Alsharski AN, Hoti Q, Usman AG (2020) Development of Computational Intelligence Algorithms for Modelling the Performance of Humanin and Its Derivatives in HPLC Optimization Method Development 9(08):110–117
- Guo F (2004) A new identification method for Wiener and Hammerstein systems. *For schungszentrum Karlsruhe*
- Huang M-J, Liaw Y-F (1995) Clinical associations between thyroid and liver diseases. *J Gastroenterol Hepatol* 10:344
- Jonklaas J (2020) Infiltration of the thyroid gland by non-thyroid malignancy: a literature review reveals this to be an unusual cause of hyperthyroidism. *J Clin Transl Endocrinol* 20(February):100221. <https://doi.org/10.1016/j.jcte.2020.100221>
- Kim S, Singh VP (2014) Modeling daily soil temperature using data-driven models and spatial distribution. *Theoret Appl Climatol* 118(3):465–479. <https://doi.org/10.1007/s00704-013-1065-z>
- Loos S, Shin CM, Sumihar J, Kim K, Cho J, Weerts A (2019) Ensemble data assimilation methods for improving river water quality forecasting accuracy. *Water Res*. <https://doi.org/10.1016/j.watres.2019.115343>
- Muhammad Ghali U, Alhosen Ali Degm M, Nouri Alsharski A, Hoti Q, Garba Usman A (nd) Development of computational intelligence algorithms for modelling the performance of humanin and its derivatives in HPLC optimization method development. www.ijstr.org
- Nourani V, Elkiran G, Abba SI (2018) Wastewater treatment plant performance analysis using artificial intelligence - An ensemble approach. *Water Sci Technol* 78(10):2064–2076. <https://doi.org/10.2166/wst.2018.477>
- Pham QB, Abba SI, Usman AG, Linh NTT, Gupta V, Malik A, Costache R, Vo ND, Tri DQ (2019) Potential of hybrid data-intelligence algorithms for multi-station modelling of rainfall. *Water Resour Manage*. <https://doi.org/10.1007/s11269-019-02408-3>
- Piantanida E, Ippolito S, Gallo D, Masiello E, Premoli P, Cusini C, Rosetti S, Sabatino J, Segato S, Trimarchi F, Bartalena L, Tanda ML (2020) The interplay between thyroid and liver: implications for clinical practice. *J Endocrinol Invest* 43(7):885–899. <https://doi.org/10.1007/s40618-020-01208-6>
- Pludowski P, Karczmarewicz E, Bayer M, Carter G, Chlebna-Sokół D, Czech-Kowalska J, Dębski R, Decsi T, Dobrzańska A, Franek E, Głuszko P, Grant WB, Holick MF, Yankovskaya L, Konstantynowicz J, Książyk JB, Książkowska-Orłowska K, Lewiński A, Litwin M, Żmijewski MA (2013) The role of EBV in thyroid disease. *Endokrynol Pol* 64(4):319–327
- Punekar P, Sharma AK, Jain A (2018) A study of thyroid dysfunction in cirrhosis of liver and correlation with severity of liver disease. *Indian J Endocrinol Metab* 22(5):645–650. https://doi.org/10.4103/ijem.IJEM_25_18

- Soltani M, Omid M, Alimardani R (2015) Egg quality prediction using dielectric and visual properties based on artificial neural network. *Food Anal Methods* 8(3):710–717. <https://doi.org/10.1007/s12161-014-9948-x>
- Usman AG, İşik S, Abba SI (2020) A novel multi-model data-driven ensemble technique for the prediction of retention factor in HPLC method development. *Chromatographia* 83(8):933–945. <https://doi.org/10.1007/s10337-020-03912-0>
- Usman AG, İşik S, Abba SI (2021) Hybrid data-intelligence algorithms for the simulation of thymoquinone in HPLC method development. *J Iran Chem Soc.* <https://doi.org/10.1007/s13738-020-02124-5>
- Vapnik V (1995) The nature of statistical learning theory, p 188. <https://doi.org/10.1007/978-1-4757-2440-0>
- Zhang Y, Wu W, Liu Y, Guan Y, Wang X, Jia L (2020) The impact of TSH levels on clinical outcomes 14 days after frozen-thawed embryo transfer, pp 1–7

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)