scientific reports

OPEN

Prediction of cell migration potential on human breast cancer cells treated with *Albizia lebbeck* ethanolic extract using extreme machine learning

Huzaifa Umar[1]✉, Maryam Rabiu Aliyu[2], Abdullahi Garba Usman[1,3], Umar Muhammad Ghali[4], Sani Isah Abba[5] & Dilber Uzun Ozsahin[6,7]✉

Cancer is one of the major causes of death in the modern world, and the incidence varies considerably based on race, ethnicity, and region. Novel cancer treatments, such as surgery and immunotherapy, are ineffective and expensive. In this situation, ion channels responsible for cell migration have appeared to be the most promising targets for cancer treatment. This research presents findings on the organic compounds present in *Albizia lebbeck* ethanolic extracts (ALEE), as well as their impact on the anti-migratory, anti-proliferative and cytotoxic potentials on MDA-MB 231 and MCF-7 human breast cancer cell lines. In addition, artificial intelligence (AI) based models, multilayer perceptron (MLP), extreme gradient boosting (XGB), and extreme learning machine (ELM) were performed to predict in vitro cancer cell migration on both cell lines, based on our experimental data. The organic compounds composition of the ALEE was studied using gas chromatography-mass spectrometry (GC–MS) analysis. Cytotoxicity, anti-proliferations, and anti-migratory activity of the extract using Tryphan Blue, MTT, and Wound Heal assay, respectively. Among the various concentrations (2.5–200 µg/mL) of the ALEE that were used in our study, 2.5–10 µg/mL revealed anti-migratory potential with increased concentrations, and they did not show any effect on the proliferation of the cells (P < 0.05; n ≥ 3). Furthermore, the three data-driven models, Multi-layer perceptron (MLP), Extreme gradient boosting (XGB), and Extreme learning machine (ELM), predict the potential migration ability of the extract on the treated cells based on our experimental data. Overall, the concentrations of the plant extract that do not affect the proliferation of the type cells used demonstrated promising effects in reducing cell migration. XGB outperformed the MLP and ELM models and increased their performance efficiency by up to 3% and 1% for MCF and 1% and 2% for MDA-MB231, respectively, in the testing phase.

Cancer is a primary cause of death in the modern world, and the incidence varies considerably based on race, ethnicity, and region. Several studies revealed metastases responsible for 90% of cancer deaths[1–4]. Breast cancer (BCa) is among the leading causes of death in the United States and other parts of the world[5]. A recent report by the WHO (2018) shows that breast cancer accounts for about 0.627 million women's mortality. However, BCa patients' mortality rate increases from systematic metastases to distant organs[5].

Therefore, there is a need for effective chemo-preventive agents with fewer side effects to suppress tumour metastasis. BCa metastasis involves complex processes established by various pathways and factors. It starts with

[1]Near East University, Operational Research Centre in Healthcare, TRNC Mersin 10, 99138 Nicosia, Turkey. [2]Department of Energy System Engineering, Cyprus International University, Northern Cyprus via Mersin 10, 99258 Nicosia, Turkey. [3]Department of Analytical Chemistry, Faculty of Pharmacy, Near East University, TRNC, Mersin 10, 99138 Nicosia, Turkey. [4]Department of Chemistry, Faculty of Natural and Applied Sciences, Firat University, Merkezi, 23199 Elazig, Turkey. [5]Interdisciplinary Research Centre for Membranes and Water Security, King Fahd University of Petroleum and Minerals, 31261 Dhahran, Saudi Arabi. [6]Department of Medical Diagnostic Imaging, College of Health Sciences, University of Sharjah, P.O. Box 27272, Sharjah, United Arab Emirates. [7]Research Institute for Medical and Health Sciences, University of Sharjah, P.O. Box 27272, Sharjah, United Arab Emirates. ✉email: Huzaifa.umar@neu.edu.tr; dozsahin@sharjah.ac.ae

cell motility in the primary affected site to other distant tissue, blood, or lymph vessels[6,7]. Metastasis involves exiting cancerous cells from the primary site to invade various organs and tissues, facilitating cellular migration, invasion, and adhesion mediated by the Ca^{2+} signalling process[8]. Metastasis maintained primary tumour-differentiated characteristics such as cell-to-cell contact, signalling, and behaviour. Transforming Growth Factor (TGF) and Epidermal Growth Factor (EGF) signalling can trigger breast cancer tumour motility[9,10]. Consequently, understanding the mechanisms behind BCa metastasis and combatting it is paramount in defeating the war against BCa[11].

Therapies and drugs are available to cure cancer, but still, there is a need for effective therapies and targeted plant-based medications with fewer side effects[12]. Various chemotherapeutic treatments with paclitaxel and anthracyclines might induce apoptosis, inhibit cell proliferation, and affect cellular activity[13]. Natural products such as plants, marine organisms, and microorganisms have revealed potential in cancer treatments[14]. Scientists continuously search for anticancer drugs that will have low toxicity and side effects but high efficacy. One such approach is targeting apoptotic cells, usually targeted as a step during therapy. Novel cancer treatments such as surgery and immunotherapy are not obligatory and quite expensive. In this situation, ion channels responsible for cell migration (metastasis) have appeared to be the most promising targets in the last two decades[14,15].

Albizia lebbeck leaves are natural products rich in flavonoids and demonstrate antitumor activity against HepG2 hepatoma cells. Zinc oxide nanoparticles synthesized using plant stem bark revealed cytotoxic activity against highly and weakly metastatic human BCa cells[15,16]. Quercetin is one such flavonoid founds in *Albizia lebbeck* that is not toxic, and it demonstrates many biological actions, including anticancer, anti-inflammatory, and antioxidant activity[4,18,19]. Quercetin has also induced cell circle arrest and apoptosis through signalling pathways and FOXO3a modification in breast cancer cells[20].

Modelling cancer cell migration in vitro has been a venerable challenge due to various complexities in cellular and molecular regulatory mechanisms in the system[21]. Conventionally, studies report the employment of rule-based greedy algorithms within agent-based modelling (ABM) in the modelling of cell migration[22,23]. Zhang et al. describe metastasis as the primary cause of death in many breast cancer patients. The cancer cell migration behaviour was analysed computationally by applying high-content imaging and microfluidic single-cell migration. Random forest decision and artificial neural network (ANN) were employed in the study. The results indicated higher accuracy regarding the prediction of cell movement[24]. Ravdin et al. described the application of artificial intelligence for predicting clinical outcomes of node-positive breast cancer patients. The status of the hormone receptor, tumour size, age of the patient, as well as relapse status were used as the input variables. The results prove a satisfactory neural network in predicting cancer status in patients[25]. Furthermore, Jerez et al. employed various machine learning and statistical methods that can be used to simulate the recurrence of breast cancer patients. The results prove the reliability of the machine learning data algorithms over the classical statistical processes in the simulation of breast cancer outcomes[26]. Based on the studies mentioned above from the technical literature, it can be observed that the applications of artificial intelligence are of paramount importance, which shows reliable and satisfactory results over the classical statistical methods. Moreover, since the developments of artificial intelligence-based models, this is the first work conducted in the technical literature depicting the application of these kinds of novel artificial intelligence models for the prediction of lateral motility in human BCa cells. However, we employed non-linear models in our study due to their flexibility, predictability, precision,and interpretability.

In this present study, the organic compounds present in *Albizia lebbeck* ethanolic extracts (ALEE), as well as their impact on anti-migratory, antiproliferative and cytotoxic potential on MDA-MB 231 and MCF-7 human BCa cell lines. In addition, the artificial intelligence (AI) based models, multilayer perceptron (MLP), extreme gradient boosting (XGB), and extreme learning machine (ELM), were performed to predict in vitro cancer cells migration on MDA-MB 231 and MCF-7cells, based on our experimental data.

Materials and methods
Plant material
Fresh stem barks of *A. lebbeck* were collected during the rainy season (April to October) from northern Nigeria, a town called Tabuli, part of Gaya Local Government, Kano State, during their flowering stage and dried at room temperature. The *A. lebbeck* stem bark collection follows all the applicable international standards, guidelines, and laws. The plant specimen was authenticated by Dr. Bala Sidi Aliyu, and deposited with voucher specimen number BUKHAN187 at the herbarium Plant Biology Department, Faculty of Science, Bayero University Kano.

Sample preparation
Dried *Albizia lebbeck* stem barks were pulverised to clear powder and subjected to flask extraction using 99.9% methanol as extraction solvent. Powdered *A. lebbeck* stem bark (50 g) was soaked in an Erlenmeyer flask containing methanol (500 mL) and placed under continual shaking for 48 at room temperature[27]. Whatman filter paper No.1 was used to filter the extract and concentrate it under reduced pressure using a Rotary evaporator. The concentrated extract was dried completely at 40 °C in an oven and stored at 4 °C before the analysis.

Phytochemical analysis of the extract
The ALEE extracts were analysed for their total flavonoid (TFC) and total phenolic content (TPC) using standard spectrophotometric methods[28,29]. To determine TFC determination, ALEE (1 mg/mL) was mixed with NaNO2 solution (5%), 10% AlCl3, and 1 M NaOH, and absorbance was measured at 510 nm. Folin-Ciocalteu reagent was added to ALEE (10:1) for TPC, followed by incubation with Na_2CO_3 (7.5%) and absorbance measurement at 760 nm. Results are presented as mg quercetin equivalent (QE)/g dry extract and gallic acid equivalents (µg GAEs/g dry extract).

We utilised gas chromatography-mass spectrometry (GC–MS) to analyse the organic composition of ALEE. We first created a crude extract in ethanol (1 mg/mL) and filtered it via a 0.22 μm syringe filter. Then, we injected it into a Shimadzu GC–MS–QP2010 plus analyser with helium as the carrier gas at a steady flow rate of 1 mL/min. The oven temperature was set at 50 °C for 2 min and gradually increased by 7 °C/min. We assessed the mass spectra at a scanning interval of 0.5 s, with a complete scan range from 25 to 1000 m/z, employing a Quadrupole mass detector. Ultimately, we identified the existing compounds by scrutinising the spectrum via the WILLEY7 MS library.

Cell models and culture conditions

MDA-MB 231 (strongly metastatic) and MCF-7 (weakly metastatic) BCa cell lines were obtained as a gift from Imperial College London (UK) and stored at the Biotechnology Research Centre (BCR) of Cyprus International University. The BCR ethical committee (BRCEC2011-01) approved using these cell lines in our study. We cultured the cells in Dulbecco's Modified Eagle's Medium (DMEM) (Gibco by Life Technology USA), supplemented with 2 mM L-glutamine, penicillin, and 10% fetal bovine serum (FBS), and maintained them in a sterile incubator at 37 °C and 5% CO_2.

Toxicity and proliferation assay

We conducted a tryphan blue dye exclusion assay, following the guidelines provided by Fraser et al.[31], to measure the level of cytotoxicity in BCa cells. We administered various doses, 0, 2.5, 10, 25, 50, 100 and 200 μg/mL, to the cells and observed them for 24, 48, and 72 h. After this period, we replaced the medium with a diluted tryphan blue solution, formulated by mixing 0.25 ml of the dye with 0.8 ml of medium. This assay accurately determined the extent of cytotoxicity present in the cells. Data are presented as averages of 3×30 measurements.

The proliferation of MDA-MB 231 (strongly metastatic) and MCF-7 (weakly metastatic) BCa cells treated with ALEE extracts were assessed using MTT (3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyltetrazolium bromide) reagent Sigma-Alderich) as described by Fraser et al. (1990) with some adjustments. BCa cells (3×10^4 cells/mL) cultured in tissue plates (12-well) were treated with 10, 5, 2.5, and 0 μg/mL of ALEE extracts and incubated for 24, 48, and 72 h. Treatments and culture medium (DMEM) were replaced every 24 h. Microplate Reader (ELX 800™) was used to measure the absorbance of the treated cell and control at 570 nm. All the experiment was performed at least thrice in triplicates ($n \geq 3$).

Wound heal assay

A wound heal assay was carried out to evaluate the anti-metastatic potential of ALEE extracts against highly metastatic (MDA-MB 231) and weakly metastatic (MCF-7) cells using the method of Fraser et al. with some modifications. Cells were plated in 35 mm culture dishes, and parallel and intersecting lines were drawn on the culture dishes[31]. Briefly, 1×10^6/mL and 5×10^5/mL cells per dish of MCF-7 and MDA-MB 231, respectively, were plated on 35 mm culture dishes, and three scratch lines were made using pipette tips (200 μL) after the cell settled. The initial and subsequent wounds causedwere captured using a camera (Leica, Germany) attached to an inverted microscope at × 100 magnification, and image processing software (ImageJ) was used to analyse the recovery wound area (cell migration) by migrating cells using Eq. (1).

$$MoI = 1 - \left(\frac{Wt}{W0}\right) \tag{1}$$

Mo I, motility index; W_t, the wound width at 24 or 48 h; W_0, initial wound width at 0 h.

Modelling approach

The study of the science of data is critical in any driven-model data-driven model. The accuracy of the data was tested using XGB, ELM, and MLP algorithms with MATLAB (R2021a). In this work, various models were proposed for the in vitro cancer metastasis prediction in MDA-MB 231 and MCF-7 cells, respectively. The data was collected from our experimental data set (n ≥ 80) to reveal the accuracy of the algorithms. In this way, two parameters were used as input variables, i.e. the motility index on the cells and the concentration of the extract, respectively. The two parameters we considered in modelling were the concentration of the extract and the motility index, although other parameters can be utilized for the same purpose. The models used have a learning algorithm with a single layer, and a fast learning rate and both the hidden biases and input layers which process and distribute data respectively, in the network are chosen randomly. However, other variables can also be used in the simulation of in vitro cancer metastasis prediction in both cell lines. In addition, models provide details on the effectiveness of the treatment, and choosing a single model that can perform best in most circumstances is difficult for the predictors, but applying various ensemble models can reveal the best models that will fit the data. Determination of cell migration potentials in breast cancer cells treated with ALEE extract using the motility index on the cells and the extract concentration as the input parameters were the main objectives of our proposed method. The proposed flowchart of the models is shown in Fig. 1.

Extreme gradient boosting (XGB)

The XGB algorithm is a commonly used model that is highly efficient with high reproducibility in analysing and modelling data using various inputs and outputs. The method was first introduced and improved by Friedman et al.[32], and it plays an essential role in the classification and regression of data. Its application in extreme learning techniques is well-known and the technique[33]. The technique uses a precise setup of up best complex decision tree algorithm to reveal good performance and speed faster than the standard gradient algorithm[34].
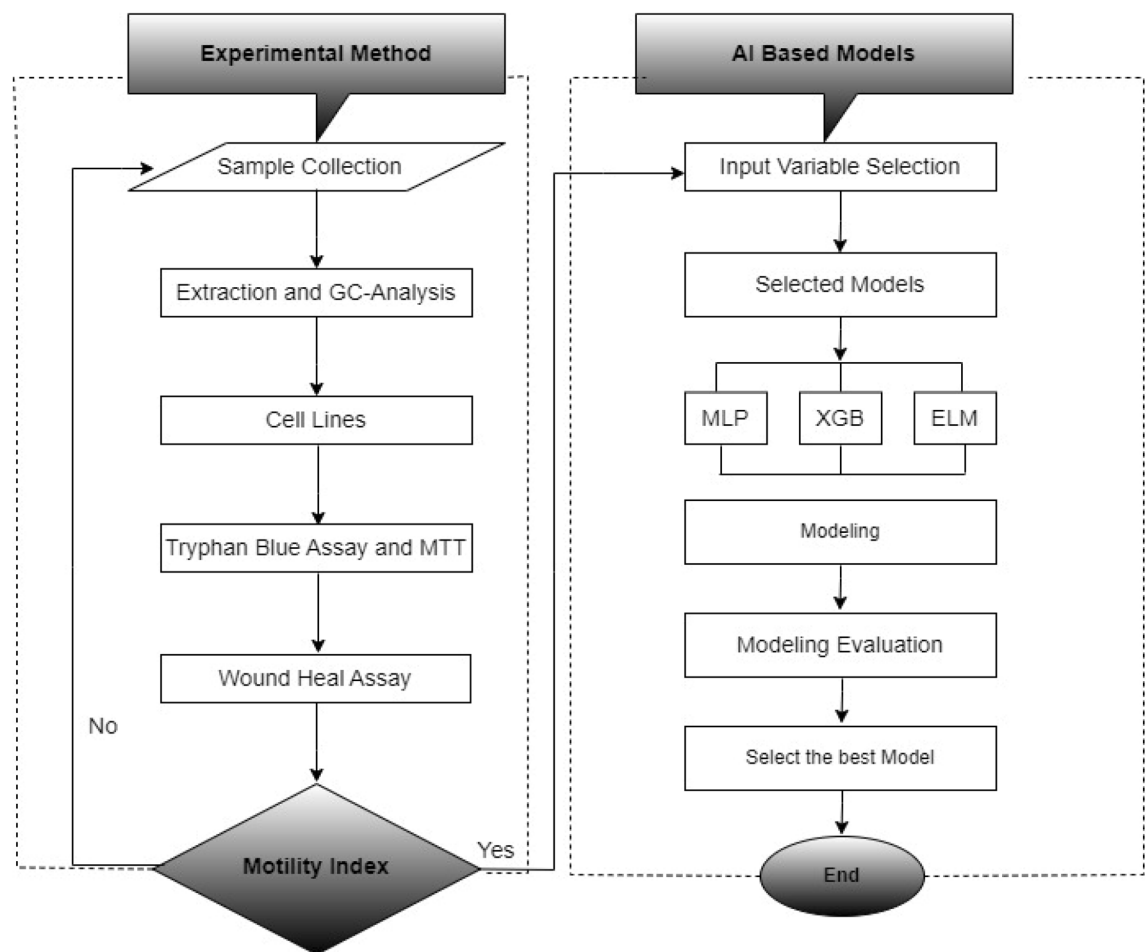
Figure 1. Proposed flowchart of experimental data-driven methods.

XGB is a machine learning ensemble technique that works similarly to Random Forest and is recognised by its classification and regression trees (CART) set. The model utilizes parallel processing to enhance learning speed, balance between variance and bias, and minimize the risk of overfitting. Furthermore, it is not the same with the decision tree (DT), whereby every leaf carries an actual score, which aids in enriching those interpretations which cannot be defined using the DT. Algorithms have been used in modelling and predicting data, and it has shown promising results. Due to this ensemble technique's wide application and excellent features, we use it to model and predict the anti-migratory potential of the cells. Given that CART $[(xi, yi) \ldots .. TK(xi, yi)]$ is the training data set of the treated cells motility index represented as x_i to predict outcomes y_i and determined using K classification, as shown in Eq. (2)[35]:

$$\hat{y} = \sum_{k=1}^{K} f_k(x_i), f_k \in F \tag{2}$$

where f_k represents independent tree structure with cells motility index scores, and F denotes the space of all CART. Optimisation of the objective is given by Eq. (3)[35]:

$$obj(\theta) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{i=1}^{t} \Omega(f_i) \tag{3}$$

The loss function is denoted l which estimates the difference between target y_i and predicted $\hat{y}_i$. The regularization function that penalises the model to avoid over-fitting is denoted as Ω, and f_i represents the simultaneous training loss function. Furthermore, the prediction value for t at step $\hat{y}_i^t$[35]:

Prediction $\hat{y}$ at the t step can be expressed as

$$\hat{y}_i^t = \sum_{k=1}^{t} f_k(x_i) = \hat{y}_i^{t-1} + f_t(x_i) \tag{4}$$

Substituting the predicted value in Eq. (4). Equation (3) can be expressed as[36]:

$$obj^t = \sum_{i=1}^{n}(y_i - (\widehat{y}_i^{t-1} + f_t(x_i)))^2 + \sum_{i=1}^{t}\Omega(f_i) \tag{5}$$

It can also be expressed as

$$obj^t = \sum_{i=1}^{n}[2\left(\widehat{y}_i^{t-1} - y_i\right)f_t(x_i) + f_t(x_i)^2 + \Omega(f_t) + constant \tag{6}$$

Looking at Taylor's expansion due to loss of function, it can be expressed in Eq. (7)[36]:

$$obj^t = \sum_{i=1}^{n}[l\left(y_i, \widehat{y}_i^{t-1}\right) + g_i f_t(x_i) + \frac{1}{2}h_i f_t(x_i)^2 + \Omega(f_t) + constant \tag{7}$$

where $g_i = \partial_{\widehat{y}_i^{t-1}}l(y_i - \widehat{y}_i^{t-1})$, and $h_i = \partial_{\widehat{y}_i^{t-1}}^2 l(y_i - \widehat{y}_i^{t-1})$. Which was described by $f_t(x) = w_{q(x)}$, and the normalised function is expressed as

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \tag{8}$$

where T represent the total number of trees, and the objective function can rewritten as

$$obj^t \approx \sum_{i=1}^{n}[g_i w_{q(x_i)} + \frac{1}{2}h_i w_{q(x_i)}^2] + \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 = \sum_{j=1}^{T}[\left(\sum_{i \in I_i} g_i\right)w_j + \frac{1}{2}\left(\sum_{i \in I_i} h_i + \lambda\right)w_j^2 + \gamma T \tag{9}$$

where $I_i = \{i|q(x_i) = j\}$ refers to the jth leaf data index. $G_j = \sum_{i \in I_i} g_i$ and $H_j = \sum_{i \in I_i} h_i$, the objective function can be written as

$$obj^t = \sum_{j=1}^{T}[G_j w_j + \frac{1}{2}\left(H_j + \lambda\right)w_j^2 + \gamma T \tag{10}$$

Performance for $q(x)$ can be achieved using the objective function and w_j, as you can see in Eqs. (11) and (12).

$$w_j^* = -\frac{G_j}{H_j + \lambda} \tag{11}$$

$$obj^* = -\frac{1}{2}\sum_{j=1}^{T}\frac{G_j}{H_j + \lambda} + \gamma T \tag{12}$$

In addition, Eq. (13) is for leaf node score during splitting, L and R are the left and right scores, and the regularisation of the additional leaf is denoted as γ.

$$Gain = \frac{1}{2}\left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda}\right] - \gamma \tag{13}$$

Extreme learning machine

The ELM model is a novel learning algorithm with a single hidden layer that works similarly to a feed-forward neural network (FNN) due to its approximation potential. And it was first introduced by Huang et al.[37]. Issues such as slower training speed and over-fitting with FNN have been addressed analytically by ELM through inversion and matrix multiplication[38]. The structure of this model contains only one layer and hidden nodes, which result in the model not requiring a learning process to calculate its parameters, and hence, it remains constant during both the training and predicting phases. In addition, ELM hidden biases and input layer are chosen randomly, and the Moore–Penrose generalised inverse function determines the output layer. The ELM revealed precision due to its robustness when applied to hydrological.

Modelling[39].

The ELM was expressed by training dataset $\{(x_1, y_1), \ldots, (x_t, y_t)\}$. Overall, the input are represented as $x_1, x_2, \ldots, x_t$ and the output as $y_1, y_2, \ldots, y_t$.

The training dataset N $(t = 1, 2, \ldots, N)$ where $x_t \in \mathbb{R}^d$ and $y_t \in \mathbb{R}$, with H hidden nodes, is given by[37] as in Eq. (14):

$$\sum_{i=1}^{H} B_i g_i(\alpha_i.x_t + \beta_i) = z_t, \tag{14}$$

Equation (14), i represents index of the hidden layer node, β_i and α_i denote the bias and weight of the random layers, and d is the number of inputs. Furthermore, the predicted weight of the output layer, model output and hidden layer neurons activation function are $B \in \mathbb{R}^H$, $Z(z_t \in \mathbb{R})$ and $G(\alpha, \beta, x)$ respectively. The best activation function is found to be the sigMoId function[40] as follows:

$$G(x) = \frac{1}{1 + exp(-x)},$$ (15)

In addition, the output layer utilizes a linear activation function, which is shown in the following equation:

$$\sum_{t=1}^{N} \|z_t - y_t\| = 0,$$ (16)

The value of B is calculated using the system of linear equations as expressed in Eq. (17) and G in Eq. (18)

$$Y = GB$$ (17)

$$G(\alpha, \beta, x) = \begin{bmatrix} g(x_1) \\ \vdots \\ g(x_N) \end{bmatrix} = \begin{bmatrix} g_1(\alpha_1.x_1 + \beta_1) & \cdots & g_L(w_H.x_1 + \beta_H) \\ \vdots & \cdots & \vdots \\ g_1(\alpha_N.x_N + \beta_1) & \cdots & g_L(w_H.x_N + \beta_F) \end{bmatrix}_{N \times H}$$ (18)

B is calculated in Eq. (19), and Y in Eq. (20).

$$B = \begin{bmatrix} B_1^T \\ \vdots \\ B_H^T \end{bmatrix}_{H \times 1}$$ (19)

$$Y = \begin{bmatrix} y_1^T \\ \vdots \\ y_N^T \end{bmatrix}_{N \times 1}$$ (20)

G is for the hidden layer. $\widehat{B}$ was calculated using "Moore–Penrose inverse function + by inverting the hidden-layer matrix" (see Eq. 21).

$$\widehat{B} = G^+ Y$$ (21)

Overall, estimated $\widehat{y}$, which denotes the predicted MoI of the cells whic,h can achieved using Eq. (22).

$$\widehat{y} = \sum_{i=1}^{H} \widehat{B}_i g_i(\alpha_i.x_t + \beta_i)$$ (22)

Multilayer perceptron

MLP, as one of the commonly applied Artificial neural networks (ANNs) composed of information processing units and an advanced simulation tool, motivated and mimicked the biological neurons. In this way, ANN, just like the human central nervous system (CNS), can solve complex problems with a non-linear and linear behaviour by combining features such as parallel processing, generalisation, learning power and decision making[41]. The general architecture of ANN consists of 3 layers with individual and different tasks: the input layer, which distributes the data in the network; the hidden layers, which process the information and the outputs, which, in addition to processing each input vector, show its work. The neurons are regarded as the smallest unit that processes the networks. The basic characteristics of MLP include using interactive connections between the neurons without advanced mathematical design to complete the information processing. Furthermore, MLP comprises input, one or more hidden and output layers in its architecture, similar to the ANN (Fig. 2)[40].

Performance objectives

To evaluate the performance efficiency of the artificial intelligence-based models used in the current study; two different metrics, where; Nash–Sutcliffe coefficient (NS) was used for understanding the fitness between the experimental and predicted values, while Root mean square error (RMSE) was used in determining the errors depicted by each model.

Hence, the Root mean square error (RMSE) was expressed as:

$$RMSE = \sqrt{\frac{1}{N} \sum_{j=1}^{N} \left((Y)_{obs,j} - (Y)_{com,j} \right)^2}$$ (23)
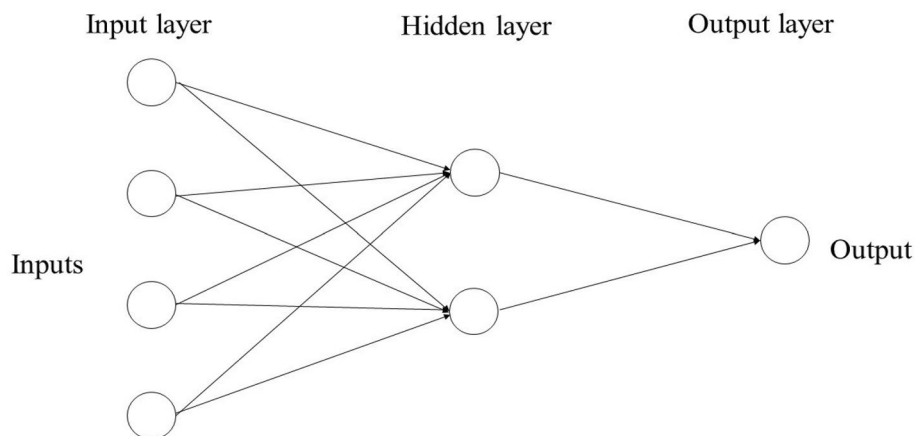
Nash–Sutcliffe coefficient (NS), expressed as:

Figure 2. Schematic diagram of MLP network structure.

$$NS = 1 - \left[\frac{\sum_{i=1}^{N} \left(Q_{obs,i} - Q_{sim,i} \right)^2}{\sum_{i=1}^{N} \left(Q_{obs,i} - \overline{Q}_{obs,i} \right)^2} \right] \infty \leq NS \leq 1 \tag{24}$$

Result and discussion
Experimental results

The study found that the ALEE contained TFC and TPC at levels of 2022.80 ± 17.83 QE μg/g and 6556.49 ± 22.52 GAE μg/g, respectively. Studies have shown that TPC is highly efficient in scavenging different oxidizing molecules, including free radicals produced during lipid peroxidation[42]. Moreover, research has revealed that flavonoids, present in various structures of phenolic compounds, possess medicinal properties. These compounds can be found in sources such as flowers, leaves, stem bark, roots, fruits and tea[43,44].

The compounds found in ALEE are listed in Table 1; their corresponding chromatogram peaks are shown in Fig. 3. We identified several significant compounds in our extract that have biological potential; some of them

S/n	RT	Peak area	Area %	Compound detected	Biological activity
1	2.00	2,506,581,156	88.55	Ethanol (CAS)	Anti-microbial[45]
2	3.08	2,369,689	0.08	(S)-(E)-(−)-4-Acetoxy-1-phenyl-2-dodecen-1-one	Anti-cancer, DNA-binding and anti-microbial[46]
3	3.77	13,527,363	0.48	Cyclotrisiloxane, hexamethyl-(CAS)	Anti-cancer, anti-oxidant and anti-microbial[47]
4	3.98	21,342,385	0.75	1,2-Dihydro-1,4-diphenylphthalazine	Anticancer[46]
5	6.46	90,085,996	3.18	Silicic acid, diethyl bis(trimethylsilyl) ester	-
6	6.74	34,182,822	1.21	1-(2-trimethylsiloxy-1,1-dideuteriovinyl)-4-trimethyl-siloxy-benzene	Anti-microbial[45]
7	7.60	47,188,509	1.67	1-cyano-5-benzoyloxy-á-ᴅ-ribofuranose	Anticancer[46]
8	11.33	40,503,790	1.43	Disiloxane, 1,3-diethoxy-1,1,3,3-tetramethyl- (CAS)	Anti-inflammatory, anti-bacterial[48]
9	13.47	3,392,203	0.12	(S)-(E)-(−)-4-Acetoxy-1-phenyl-2-dodecen-1-one	Anti-cancer, DNA-binding and anti-microbial[46]
10	13.76	15,872,886	0.56	Silicic acid, diethyl bis(trimethylsilyl) ester	-
11	15.69	5,409,162	0.19	Silicic acid, diethyl bis(trimethylsilyl) ester	-
12	16.67	8,613,439	0.30	2-(5-acetyl-2-thienyl)-1,4-naphthoquinone	Anti-microbial, anti-tumor and antiproliferative[46]
13	16.84	11,375,011	0.40	4H-3-(p-methylanilino)1-benzothiopyran-4-one 1-oxide	Anti-cancer[46]
14	17.29	3,343,231	0.12	Heptasiloxane,1,1,3,3,5,5,7,7,9,9,11,11,13,13-tetradeca-methyl-	Anti-oxidant and anti-inflammatory[48]
15	18.70	14,832,408	0.52	3-(4'-Methoxyphenyl)-1-acetyl-2-phenylindolizine	-
16	23.31	3,451,675	0.12	Cyclooctasiloxane, hexadecamethyl-	Anti-cancer, anti-oxidant and anti-microbial[47]
17	25.33	6,562,803	0.23	4-Ethylbenzoic acid, 2-butyl ester	Anti-cancer and Anti-mibroabial[47]
18	29.18	2,201,203	0.08	Cyclononasiloxane, octadecamethyl-	Anti-cancer, anti-oxidant and anti-microbial[48]
18	29.18	2,201,203	0.08	Cyclononasiloxane, octadecamethyl-	Anti-cancer, anti-oxidant and anti-microbial[46]

Table 1. Bioactive compounds identified from ALEE by GC–MS analysis and their biological activities.
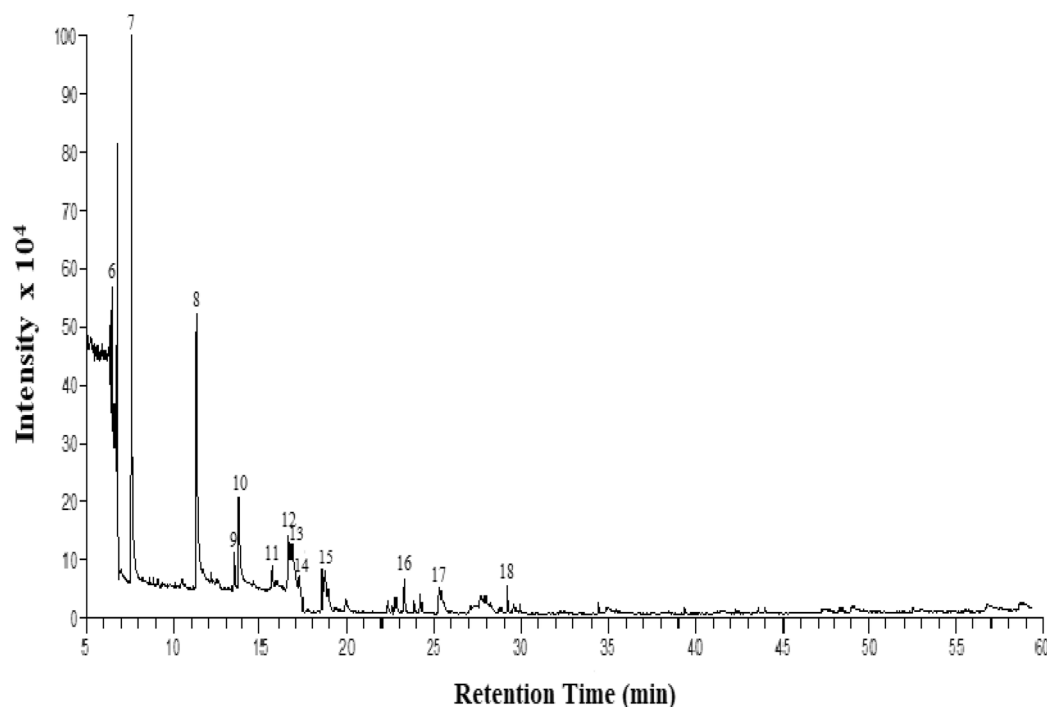
Figure 3. GC–MS chromatogram of ALEE bioactive compounds. The numbered peaks correspond to the numbers and molecules in Table 1.

are Ethanol (88.55%), Silicic acid, diethyl bis(trimethylsilyl) ester (3.18%), 1-cyano-5-benzoyloxy-á-D-ribo-furanose (1.67%), 1-(2-trimethylsiloxy-1,1-dideuteriovinyl)-4-trimethyl siloxy-benzene (1.21%), Disiloxane, 1,3-diethoxy-1,1,3,3-tetramethyl- (1.43%), 1,2-Dihydro-1,4-diphenylphthalazine (0.75%), 3-(4'-Methoxyphenyl)-1-acetyl-2-phenylindolizine (0.52%), 4H-3-(p-methylamino)1-benzothiopyran-4-one 1-oxide (0.40%). The high percentage of ethanol might be from the extraction solvent, indicating that it is not a suitable solvent for *A. lebbeck* extraction. Furthermore, the extract's anti-proliferative and anti-migratory potential might result from the phytochemicals present in the extract, and ALEE could be a good cause that will prevent the metastasis of breast cancer.

The effect of various concentrations (2.5–200 μg/mL) of ALEE on human BCa cells for 24 h and 48 h and cytotoxicity and effect on proliferation were determined using tryphan blue assay and MTT, respectively (Figs. 4 and 5). Various ALEE concentrations used in the study are 2.5, 5 and 10 μg/mL, and they demonstrated no effect on the viability of both cells. Still, treatment with a concentration between 25 and 200 μg/mL revealed significant changes ($P < 0.05$). Treatment of MDA-MB 231 with 2.5, 5 and 10 μg/mL ALEE did not significantly alter cell viability compared to untreated cells (control). Similarly, treatment of MCF-7 with 2.5, 5 and 10 μg/mL ALEE did not show significant changes compared with the control ($P > 0.05$). Studies revealed in vitro anti-proliferative potential of Silicic acid, diethyl bis(trimethylsilyl) ester separated from *Lorabthus parasiticus* on breast cancer in a dose-dependent manner, which is in agreement with our findings[49]. (S)-(E)-(−)-4-Acetoxy-1-phenyl-2-dodecen-1-one (Quercetin) isolated from green tea revealed anti-proliferative effect against in PC-3 and LNCaP human prostate cancer cells[50]. Furthermore, studies revealed that quercetin isolated from plants inhibits proliferation, signal transduction and metastasis in cancer cell lines[51].

According to the study, the concentration of ALEE did not have a notable impact on the viability and growth of MDA-MB 231 and MCF-7 human BCa cells when compared to the control group. Nonetheless, the anti-migratory capacity of the cells was examined through the wound healing assay, and it was discovered that the lateral motility index (MOI) of MDA-MB 231 decreased with an increase in ALEE concentration and incubation duration. Figure 6 indicates that 10 μg/mL of ALEE had the most optimal motility index among the other concentrations. The MOI is more in MDA-MB 231 because the cells are metastatic and aggressive. In addition, all ALEE concentrations revealed significant differences relative to the control ($P < 0.05$) (Fig. 6). Similarly, the MCF-7 MOI was reduced with increased ALEE concentration and incubation period, as shown in Fig. 6d,e, and 10 μg/mL revealed the lowest and best MOI when compared with the remaining ALEE concentrations. MCF-7 is a less aggressive and weakly metastatic cell, which could be the reason for the lower MoI compared with MDA-MB 231 cells. The ALEE concentrations (2.5–10 μg/mL) revealed a decrease in the MOI of MCF-7 cells with increased concentrations and incubation time, and 10 μg/mL revealed more effect on lateral motility followed by 5 μg/mL (Fig. 6d,e; $n \geq 3$). Nanoparticles synthesised using quercetin reached plant (*Ficus ingens*) revealed an effect on lateral motility of MDA-MB 231[13]. Medicinal plants containing quercetin as an active ingredient showed anti-metastatic activity on strongly and weakly metastatic MatLYLu and AT-2 rat prostate cancer cell models, respectively[52].
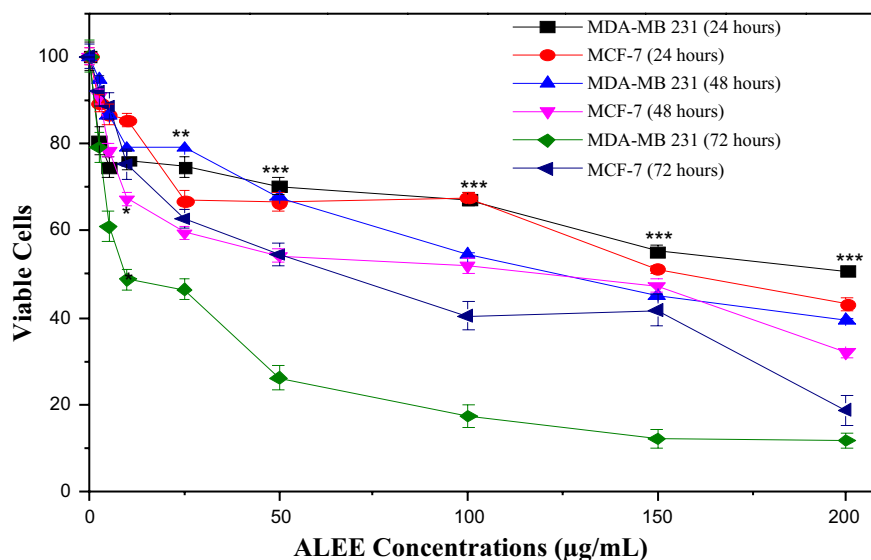
Figure 4. Cytotoxic Effect of various concentrations of ALEE MDA-MB 231 and MCF-7 breast cancer cell lines carried out using tryphan blue exclusion assay. The observations were made after 24, 48 and 72 h' incubation period. The findings are presented as mean ± SEM of at least replicates experiments ($n \geq 3$) and analysed using one-way ANOVA followed by Newman–Keuls post hoc analysis. (*) $P < 0.05$; (**) $P < 0.01$ and (***) $P < 0.0001$.

Anti-migratory potential prediction models

The AI-based models (MLP, XGB, and ELM) were analysed to predict in vitro cancer migration prediction in cells treated with ALEE based on our experimental data. Before the model calibration, statistical data analysis was conducted, as shown in Table 2. Generally, statistical analysis is done to understand the dataset. Furthermore, The AI-based models (MLP, XGB, and ELM) were analysed to predict in vitro cancer migration prediction in the MDA-MB 231 and MCF-7 human BCa, treated with ALEE based on our experimental data. The performance evaluation is checked by applying various criteria to compare the simulated and the observed values. The distribution between the different multiple parameters and the dataset used in the study was expressed as a visualised pie chart in Fig. 7, and the data set is well distributed. Furthermore, the correlation matrix shows the correlation between different parameters in a linear form. It can be seen from Fig. 8 that there is a high correlation between all the parameters, whereby the highest correlation in this study is between MDA-MB231 and MCF-7 having R-value = 0.98, and the lowest correlation exists between concentration and MCF-7 with R = 0.75. Similarly, the correlation matrix shows a robust correlation between all the variables and is in conformity with the correlation revealed by Adun et al.[53].

The modelling performance of MLP, XGB, and ELM models, treated MDA-MB 231 and MCF-7, were compared to each other using RMSE and NSE, as shown in Table 2. Based on the predictive comparison of the models in Table 3, it can be shown clearly that all three data-driven models (MLP, XGB and ELM) can simulate the in vitro cancer migration potential prediction in the human BCa cells. XGB depicted the superiority over the other two non-linear models in the testing and training stages for modelling the performance of the cells. In regards to their error values, XGB shows the lowest RMSE values, XGB-MCF-7 = 0.0039 and XGB-MDAMB231 = 0.0025 in the testing phase, and the NSE as a goodness of fits which shows that XGB equally outperformed all the other AI-based models MLP and ELM and increase their performance efficiency up to 3% and 1% for MCF and 1% and 2% for MDA-MB231 respectively in the testing phase. The relative predictive accuracy regarding the relative error can also be demonstrated using a bar chart (Fig. 9), which reflects the performance of in vitro cancer metastasis prediction in human BCa cells in a surface radar chart showing the scale of NSE in the training and testing phases. It has been reported that the radar scale generally ranges between 0 and 1. The radar chart performance demonstrated that the versions in terms of NSE of treated BCa cell migration in highly and weakly metastatic human BCa cell lines follow the following order: XGB > ELM > MLP for MCF-7, and XGB > MLP > ELM for MDA-MB 231, respectively (Fig. 10). BCa subtypes were identified based on the immune signature in the tumour microenvironment for accurate assessment and treatment of BCa using the MLP model, and the study outcomes conform with our results[54]. The metastatic status of BCa and new therapeutic target provision were predicted using an efficient XGB model optimized by a grid search algorithm[55]. In addition, Benign or malignant types of BCa were classified using classification robustness ELM and based on
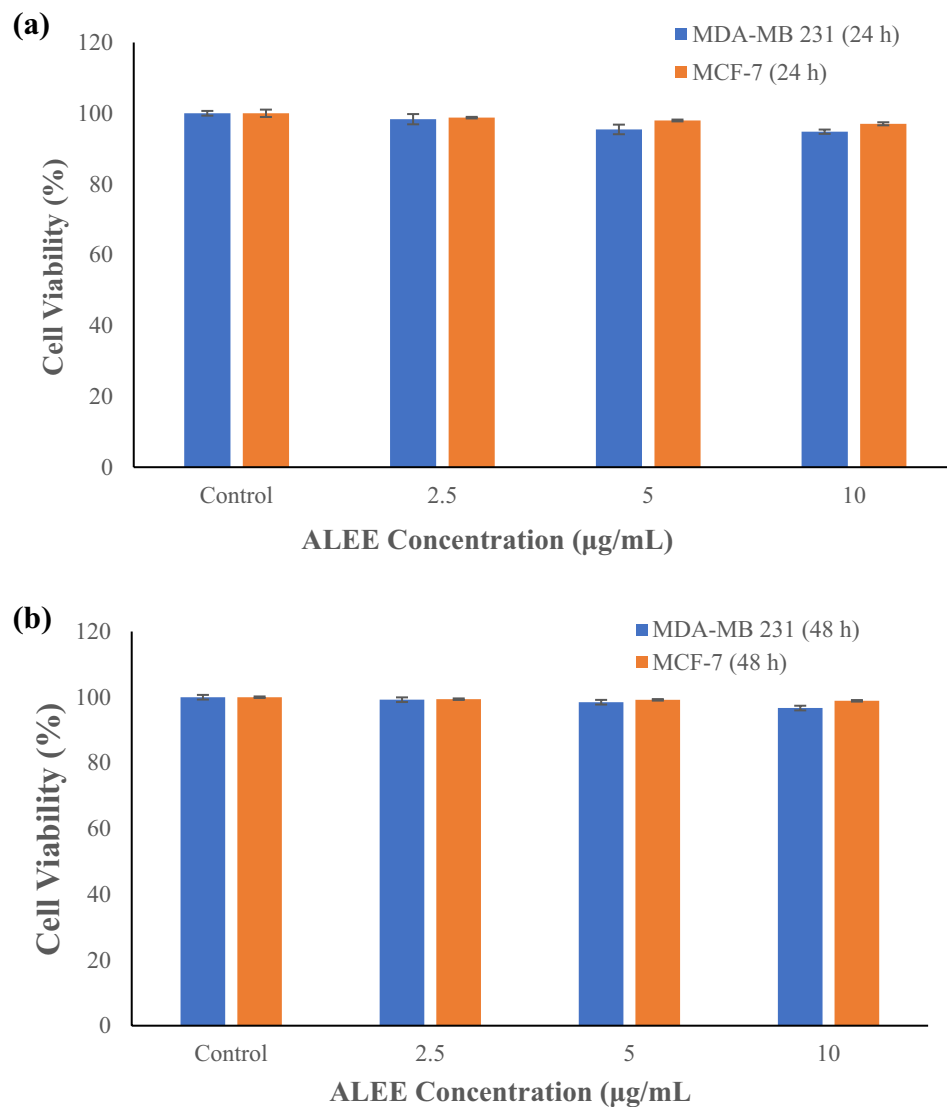
(a)

(b)

Figure 5. Effect of ALEE on proliferation of MDA-MB 231 and MCF-7. MDA-MB 231 and MCF-7 were treated with various concentrations of ALEE for (**a**) 24 h and (**b**) 48 h carried out using MTT Assay. Data are presented as mean ± SEM of at least replicates experiments ($n \geq 3$), and analysed using one-way ANOVA followed by Newman-Keuls post hoc analysis. (*) $P < 0.05$; (**) $P < 0.01$ and (***) $P < 0.0001$.

input mammograms, and the outcomes are in agreement with our findings[56]. Similarly, the methanolic extract of *A. lebbeck* demonstrated good performance using other AI-based models[57].

Conclusion

Our study has uncovered promising organic compounds in ALEE that possess medicinal properties, potentially aiding in the prevention of metastasis in human breast cancer. Interestingly, we observed that varied concentrations of the plant extract were non-toxic and had no impact on cell proliferation but displayed significant anti-migratory potential in both MDA-MB 231 and MCF-7 cells, with increasing concentration. Furthermore, we found that AI models, including MLP, XGB, and ELM, were effective in predicting the anti-migratory potential of ALEE. XGB demonstrated the highest performance efficiency, outperforming MLP and ELM models by 3% and 1% for MCF and 1% and 2% for MDA-MB231 during the testing phase. However, further studies are required to ascertain the anti-metastatic potential of the plant using various cell lines as well as to validate the anti-migratory potential of this plant, and additional computational models should be employed to improve performance.
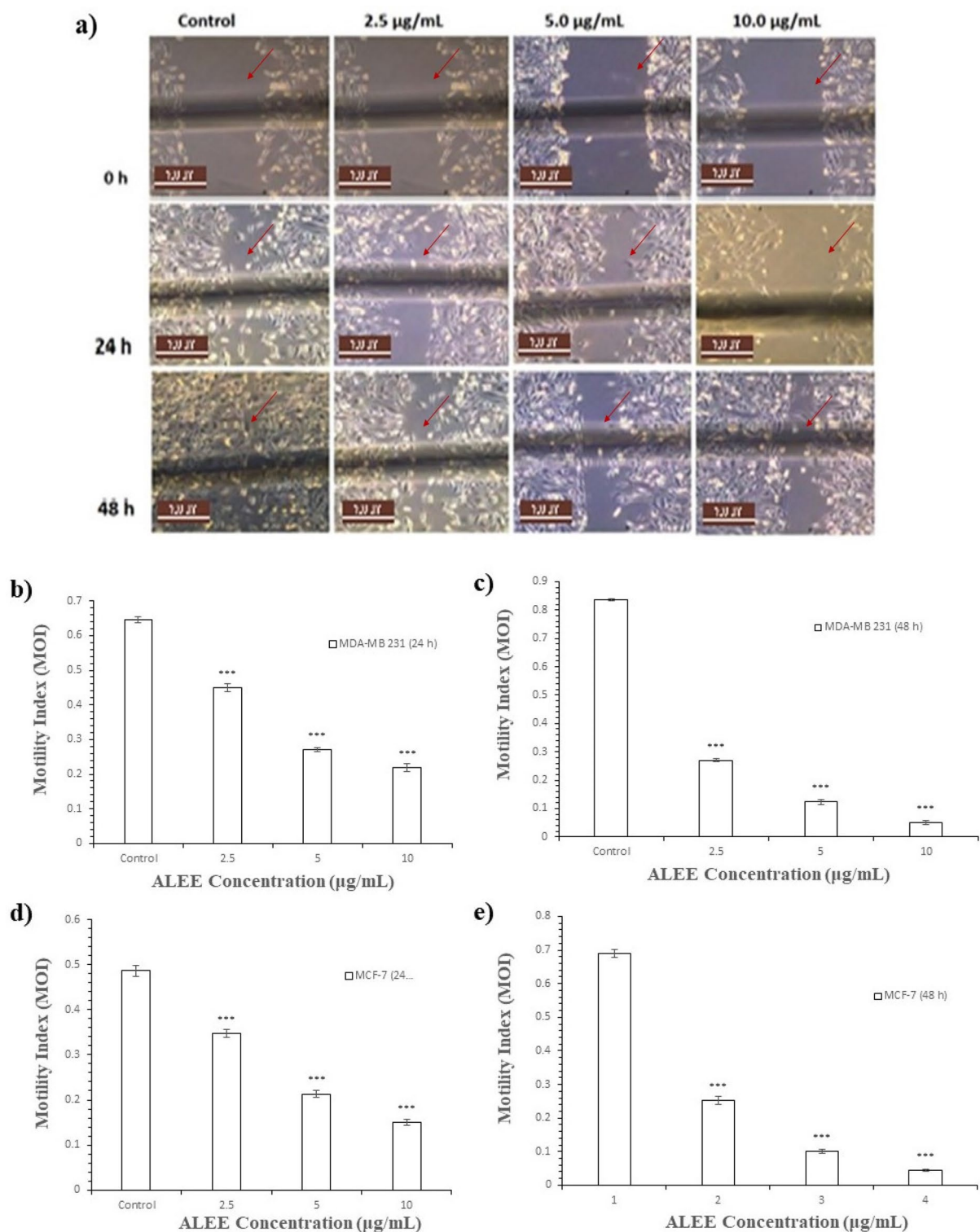
Figure 6. Effect of ALEE on lateral motility of MDA-MB 231 and MCF-7. (**a**) Typical phase-contrast light-microscopy (×10) images obtained from wound healing assays of MDA-MB 231 cell following 24 h and 48 h incubation to show the wound. Scale bar (100 μm) applicable to all panels. Bar diagram showing motility index of MDAMB 231 cell following (**b**) 24 h incubation ± ALEE (**c**) 48 h incubation ± ALEE. Bar diagram showing motility index of MCF-7 cell following (**d**) 24 h incubation ± ALEE (**e**) 48 h incubation ± ALEE Data are presented as mean ± SEM of at least replicates experiments ($n \geq 3$).

Variables	Mean	SE	Median	SD	Kurtosis	Skewness	Min	Max
Conc	4.375	0.48138	3.75	3.728753	− 1.14951	0.445876	0	10
MCF-7	0.405333	0.017451	0.47	0.135177	− 0.5607	− 1.10911	0.12	0.54
MDA-MB 231	0.402167	0.010146	0.435	0.078591	− 0.5463	− 1.04085	0.24	0.49

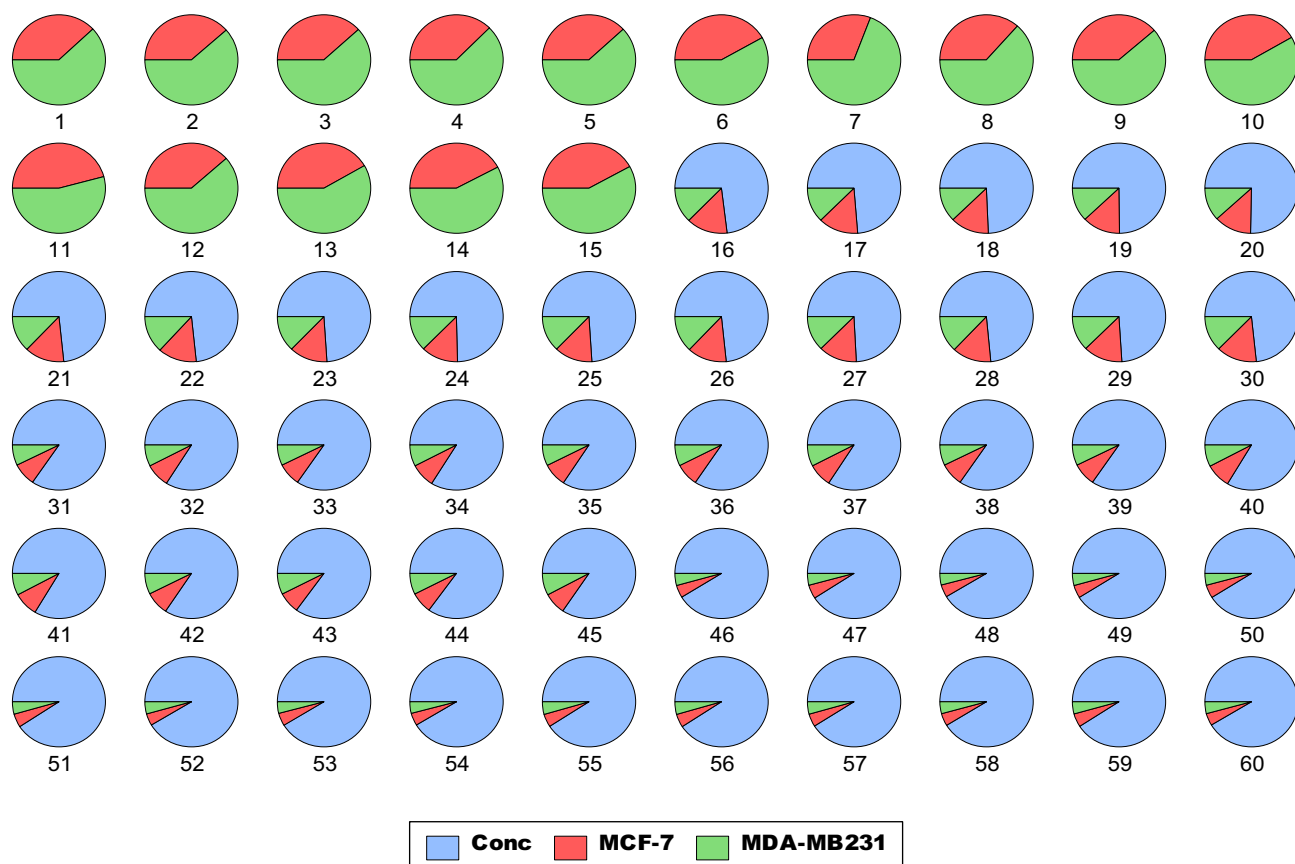**Table 2.** Descriptive statistic. *Conc.* concentration, *SE* standard error of mean, *SD* standard deviation, *Min.* minimum, *Max.* maximum.

Figure 7. Visualized pie-chart distribution of our data set for both MDA-MB 231 and MCF-7 cells treated with various concentration of ALEE extract.
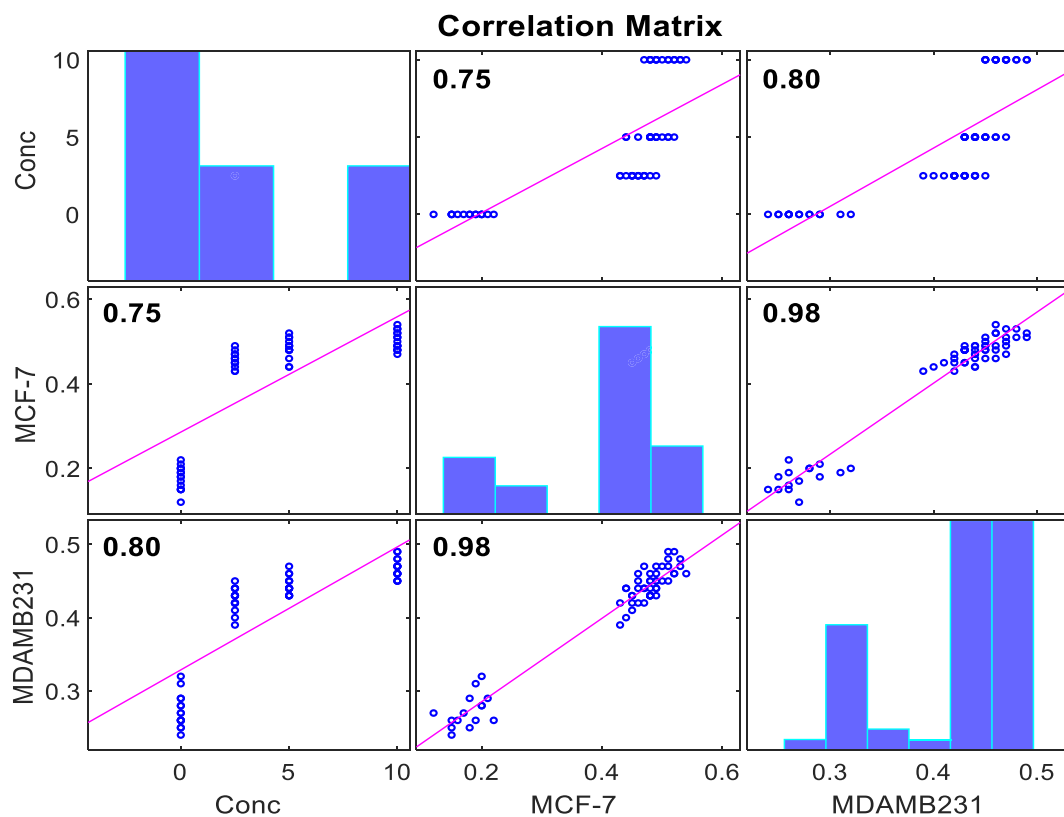
Figure 8. Correlation matrix between the raw data.

	Calibration		Verification	
	NSE	RMSE	NSE	RMSE
MLP-MCF-7	0.9964	0.0263	0.9692	0.0302
XGB-MCF-7	0.9999	0.0033	0.9975	0.0039
ELM-MCF-7	0.9998	0.0052	0.9974	0.0064
MLP-MDA-MB231	0.9955	0.0181	0.9871	0.0356
XGB-MDA-MB231	0.9985	0.0018	0.9974	0.0025
ELM-MDA-MB231	0.9924	0.0232	0.9728	0.0432

Table 3. Result of the ANN, ANFIS and MLR models.
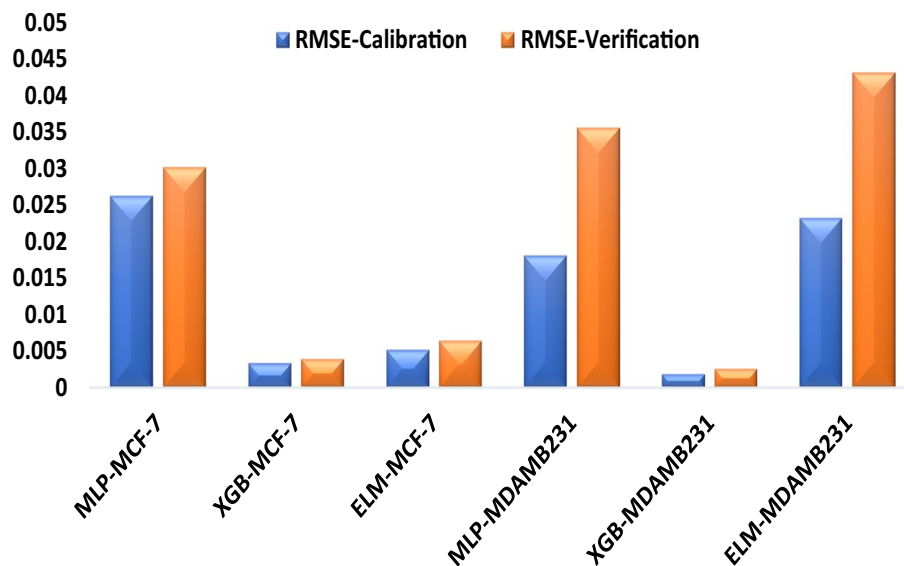
Figure 9. Relative mean square error (RMSE) for predicting MDA-MB 231 and MCF-7 human Breast Cancer treated with ALEE in both the calibration and verification stages.
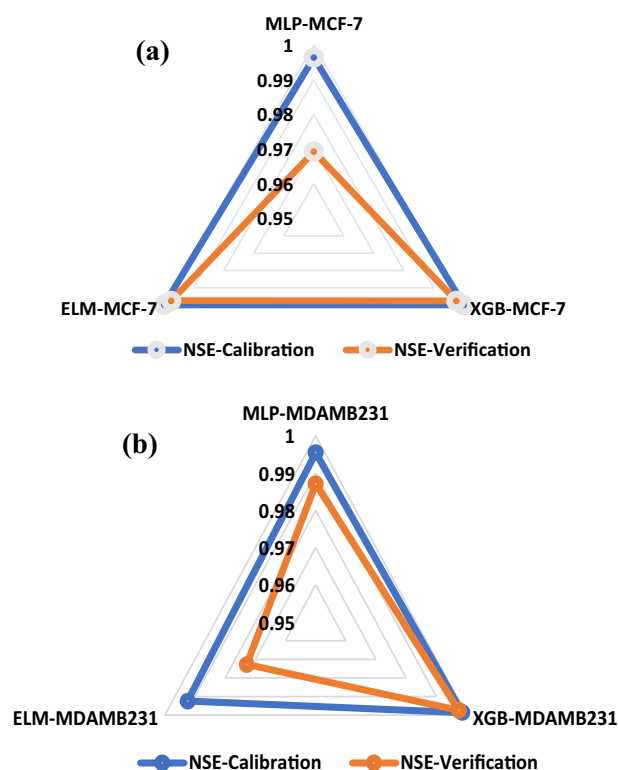

Figure 10. Radar chart for various variation determination and correlation coefficients (**a**) MCF-7 (**b**) MDA-MB 231.

Data availability
All data is included in the manuscript.

References

1. Bogenrieder, T. & Herlyn, M. Axis of evil: Molecular mechanisms of cancer metastasis. *Oncogene* https://doi.org/10.1038/sj.onc.1206757 (2003).
2. Dillekås, H., Rogers, M. S. & Straume, O. Are 90% of deaths from cancer caused by metastases?. *Cancer Med.* https://doi.org/10.1002/cam4.2474 (2019).
3. Geiger, T. R. & Peeper, D. S. Metastasis mechanisms. *Biochim. Biophys. Acta Rev. Cancer* https://doi.org/10.1016/j.bbcan.2009.07.006 (2009).
4. Gupta, G. P. & Massagué, J. Cancer metastasis: Building a framework. *Cell* https://doi.org/10.1016/j.cell.2006.11.001 (2006).
5. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2019. *CA Cancer J. Clin.* https://doi.org/10.3322/caac.21551 (2019).
6. Sahai, E. Illuminating the metastatic process. *Nat. Rev. Cancer* https://doi.org/10.1038/nrc2229 (2007).
7. Weigelt, B. & Peterse, J. L. Breast cancer metastasis: Markers and models. *Nat. Rev. Cancer* https://doi.org/10.1038/nrc1670 (2005).
8. Penna, A. *et al.* PI3-kinase promotes TRPV2 activity independently of channel translocation to the plasma membrane. *Cell Calcium* https://doi.org/10.1016/j.ceca.2006.01.009 (2006).
9. Jakowlew, S. B. Transforming growth factor-β in cancer and metastasis. *Cancer Metastasis Rev.* https://doi.org/10.1007/s10555-006-9006-2 (2006).
10. Lesko, E. & Majka, M. The biological role of HGF-MET axis in tumor growth and development of metastasis. *Front. Biosci.* https://doi.org/10.2741/2760 (2008).
11. Giampieri, S. *et al.* Localised and reversible TGFβ signalling switches breast cancer cells from cohesive to single cell motility. *Nat. Cell Biol.* https://doi.org/10.1038/ncb1973 (2009).
12. Singh, J., Hussain, Y., Luqman, S. & Meena, A. Targeting Ca2+ signalling through phytomolecules to combat cancer. *Pharmacol. Res.* https://doi.org/10.1016/j.phrs.2019.104282 (2019).
13. Kavaz, D., Umar, H. & Shehu, S. Synthesis, characterization, antimicrobial and antimetastatic activity of silver nanoparticles synthesized from Ficus ingens leaf. *Artif. Cells Nanomed. Biotechnol.* https://doi.org/10.1080/21691401.2018.1536060 (2018).
14. Leanza, L., Managò, A., Zoratti, M., Gulbins, E., & Szabo, I. Pharmacological targeting of ion channels for cancer therapy: In vivo evidences. *Biochim. Biophys. Acta Mol. Cell Res.* (2016).
15. Djamgoz, M. B. A., Coombes, R. C. & Schwab, A. Ion transport and cancer: From initiation to metastasis. *Philos. Trans. R. Soc. B Biol. Sci.* https://doi.org/10.1098/rstb.2013.0092 (2014).
16. Lam, S. K. & Ng, T. B. First report of an anti-tumor, anti-fungal, anti-yeast and anti-bacterial hemolysin from Albizia lebbeck seeds. *Phytomedicine* https://doi.org/10.1016/j.phymed.2010.08.009 (2011).
17. Umar, H., Kavaz, D. & Rizaner, N. Biosynthesis of zinc oxide nanoparticles using albizia lebbeck stem bark, and evaluation of its antimicrobial, antioxidant, and cytotoxic activities on human breast cancer cell lines. *Int. J. Nanomed.* **14**, 87–100. https://doi.org/10.2147/IJN.S186888 (2019).
18. da Silva, A. B., Cerqueira Coelho, P. L., das Neves Oliveira, M., Oliveira, J. L., Oliveira Amparo, J. A., da Silva, K. C., Costa, S. L. The flavonoid rutin and its aglycone quercetin modulate the microglia inflammatory profile improving antiglioma activity. *Brain Behav. Immunity.* https://doi.org/10.1016/j.bbi.2019.05.003 (2020).
19. Morel, I., Lescoat, G., Cogrel, P., Sergent, O., Pasdeloup, N., Brissot, P., Cillard, J. Antioxidant and iron-chelating activities of the flavonoids catechin, quercetin and diosmetin on iron-loaded rat hepatocyte cultures. *Biochem. Pharmacol.* (1993).
20. Nguyen, L. T., Lee, Y. H., Sharma, A. R., Park, J. B., Jagga, S., Sharma, G., Nam, J. S. Quercetin induces apoptosis and cell cycle arrest in triple-negative breast cancer cells through modulation of Foxo3a activity. *Korean J. Physiol. Pharmacol.* https://doi.org/10.4196/kjpp.2017.21.2.205 (2017).
21. Kusumoto, D. & Yuasa, S. The application of convolutional neural network to stem cell biology. *Inflamm. Regen.* https://doi.org/10.1186/s41232-019-0103-3 (2019).
22. Setty, Y., Cohen, I. R., Dor, Y. & Harel, D. Four-dimensional realistic modelling of pancreatic organogenesis. *Proc. Natl. Acad. Sci. USA* https://doi.org/10.1073/pnas.0808725105 (2008).
23. Wang, Z. *et al.* An observation-driven agent-based modelling and analysis framework for C. elegans embryogenesis. *PLoS ONE* https://doi.org/10.1371/journal.pone.0166551 (2016).
24. Zhang, Z. *et al.* Integrative Biology Morphology-based prediction of cancer cell migration using an artificial neural network and a random decision forest. *Integr. Biol.* **10**, 758–767. https://doi.org/10.1039/c8ib00106e (2018).
25. Ravdin, P. M., Clark, G. M., Hilsenbeck, S. G., Owens, M. A., Vendely, P., & Mcguire, W. L. *A demonstration that breast cancer recurrence can be predicted by neural network analysis.* 47–53 (1992).
26. Jerez, M. *et al. Artificial Intelligence in Medicine Missing data imputation using statistical and machine learning methods in a real breast cancer problem.* **50**, 105–115. https://doi.org/10.1016/j.artmed.2010.05.002 (2010).
27. Umar, H., Kavaz, D., Abubakar, A. L., Aliyu, M. R. & Rizaner, N. Synthesis of zinc oxide nanoparticles using *Ficus thonningii* aqueous extract and evaluation of its anti-oxidant and anti-microbial activities. *Bulgar. Chem. Commun.* **277**, 1 (2022).
28. Singleton, V. L. & Rossi, J. A. Colorimetry of total phenolics with phosphomolybdic-phosphotungstic acid reagents. *Am. J. Enol. Viticult.* **16**(3), 144–158 (1965).
29. Meda, A., Lamien, C. E., Romito, M., Millogo, J. & Nacoulma, O. G. Determination of the total phenolic, flavonoid, and proline contents in Burkina Fasan honey, as well as their radical scavenging activity. *Food Chemistry* **91**(3), 571–577 (2005).
30. Kumar, J. *et al.* Chemical composition and biological activities of trans-himalayan alga spirogyra porticalis (Muell) cleve. *Plos One.* **10**(2), 1 (2015).
31. Fraser, S. P. *et al.* Contribution of functional voltage-gated Na+ channel expression to cell behaviors involved in the metastatic cascade. *J. Cell Physiol.* **195**(3), 479–487 (2003).
32. Friedman, J. H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **1**, 1 (2001).
33. Pradhan, B., & Sameen, M. I. Predicting injury severity of road traffic accidents using a hybrid extreme gradient boosting and deep neural network approach. In *Laser Scanning Systems in Highway and Safety Assessment* (pp. 119–127) (Springer, 2020).
34. Chen, T., & Guestrin, C. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining,* 785–794. ACM (2016).
35. Chen, Z.-Y. *et al.* Extreme gradient boosting model to estimate PM2.5 concentrations with missing-filled satellite data in China. *Atmos. Environ.* **202**, 180–189. https://doi.org/10.1016/J.ATMOSENV.2019.01.027 (2019).
36. Chen, T., He, T., Benesty, M., Khotilovich, V., & Tang, Y. Xgboost: extreme gradient boosting. *R Package Version 0.4-2,* 1–4 (2015).
37. Huang, G.-B., Zhu, Q.-Y. & Siew, C.-K. Extreme learning machine: Theory and applications. *Neurocomputing* **70**(1–3), 489–501. https://doi.org/10.1016/j.neucom.2005.12.126 (2006).
38. Huang, G., Huang, G. B., Song, S. & You, K. Trends in extreme learning machines: A review. *Neural Netw.* **61**, 32–48. https://doi.org/10.1016/j.neunet.2014.10.001 (2015).
39. Yaseen, Zaher M, Allawi, M. F., Yousif, A. A., Jaafar, O., Hamzah, F. M., & El-Shafie, A. (2016). Non-tuned machine learning approach for hydrological time series forecasting. *Neural Comput. Appl.* 1–13. https://doi.org/10.1007/s00521-016-2763-0
40. Fijani, E., Barzegar, R., Deo, R., Tziritis, E. & Konstantinos, S. Design and implementation of a hybrid model based on two-layer decomposition method coupled with extreme learning machines to support real-time environmental monitoring of water quality parameters. *Sci. Total Environ.* **648**, 839–853. https://doi.org/10.1016/j.scitotenv.2018.08.221 (2019).

41. Ahmad, M. H., Usman, A. G. & Abba, S. I. Comparative performance of extreme learning machine and Hammerstein-Weiner models for modelling the intestinal hyper-motility and secretory inhibitory effects of methanolic leaf extract of *Combretumhypopilinum* Diels (Combretaceae). *In Silico Pharmacol.* **9**, 31. https://doi.org/10.1007/s40203-021-00090-1 (2021).
42. Aryal, S. *et al.* Total phenolic content, flavonoid content and antioxidant potential of wild vegetables from Western Nepal. *Plants (Basel).* **8**(4), 96. https://doi.org/10.3390/plants8040096 (2019).
43. Ayala, A., Muñoz, M. F. & Argüelles, S. Lipid peroxidation: production, metabolism, and signalling mechanisms of malondialdehyde and 4-hydroxy-2-nonenal. *Oxid. Med. Cell Longev.* **2014**, 360438. https://doi.org/10.1155/2014/360438 (2014).
44. Tungmunnithum, D., Thongboonyou, A., Pholboon, A. & Yangsabai, A. Flavonoids and other phenolic compounds from medicinal plants for pharmaceutical and medical aspects: An overview. *Medicines (Basel).* **5**(3), 93. https://doi.org/10.3390/medicines5030093 (2018).
45. Kampf, G. & Hollingsworth, A. Comprehensive bactericidal activity of an ethanol-based hand gel in 15 seconds. *Ann. Clin. Microbiol. Antimicrob.* **7**, 2. https://doi.org/10.1186/1476-0711-7-2 (2008).
46. Musa, M. A. *et al.* Cytotoxic activity of new acetoxycoumarin derivatives in cancer cell lines. *Anticancer Res.* **31**(6), 2017–2022 (2011).
47. Shin, S. Y. *et al.* Anticancer activities of cyclohexenone derivatives. *Appl. Biol. Chem.* **63**, 82. https://doi.org/10.1186/s13765-020-00567-1 (2020).
48. Ezez, D., Mekonnen, N. & Tefera, M. Phytochemical analysis of *Withania somnifera* leaf extracts by GC-MS and evaluating antioxidants and antibacterial activities. *Int. J. Food Prop.* **26**(1), 581–590. https://doi.org/10.1080/10942912.2023.2173229 (2023).
49. Devi, Y. P., Uma, A., Narasu, M. L. & Kalyani, C. Anticancer activity of gallic acid on cancer cell lines, HCT15 and MDA MB 231. *Int. J. Res. Appl. Nat. Soc. Sci.* **1**, 1 (2014).
50. Wang, P., Heber, D. & Henning, S. M. Quercetin increased the antiproliferative activity of green tea polyphenol (−)-epigallocatechin gallate in prostate cancer cells. *Nutr. Cancer* https://doi.org/10.1080/01635581.2012.661514 (2012).
51. Vidya Priyadarsini, R. *et al.* The flavonoid quercetin induces cell cycle arrest and mitochondria-mediated apoptosis in human cervical cancer (HeLa) cells through p53 induction and NF-κB inhibition. *Eur. J. Pharmacol.* **1**, 1 (2010).
52. Gumushan-Aktas, H. & Altun, S. Effects of Hedera helix L. extracts on rat prostate cancer cell proliferation and motility. *Oncol. Lett* https://doi.org/10.3892/ol.2016.4941 (2016).
53. Adun, H., Kavaz, D., Dagbasi, M., Umar, H. & Wole-Osho, I. An experimental investigation of thermal conductivity and dynamic viscosity of Al2O3-ZnO-Fe3O4 ternary hybrid nanofluid and development of machine learning model. *Powder Technol.* **394**, 1121–1140 (2021).
54. Yang, X. *et al.* Immune subtype identification and multi-layer perceptron classifier construction for breast cancer. *Front Oncol.* **12**, 943874. https://doi.org/10.3389/fonc.2022.943874 (2022).
55. Li, Q. *et al.* XGBoost-based and tumor-immune characterized gene signature for the prediction of metastatic status in breast cancer. *J. Transl. Med.* https://doi.org/10.1186/s12967-022-03369-9 (2022).
56. Sannasi Chakravarthy, S. R. & Rajaguru, H. Automatic detection and classification of mammograms using improved extreme learning machine with deep learning. *IRBM* **43**(1), 49–61. https://doi.org/10.1016/j.irbm.2020.12.004 (2022).
57. Umar, H. *et al.* Prediction of cell migration in MDA-MB 231 and MCF-7 human breast cancer cells treated with *Albizia Lebbeck* Methanolic extract using multilinear regression and artificial intelligence-based models. *Pharmaceuticals* **16**, 858. https://doi.org/10.3390/ph16060858 (2023).

Author contributions

H.U.: conceptualization, software, validation, formal analysis, writing—review and editing, supervision. M.R.A.: conceptualization, formal analysis, supervision. U.M.G.: conceptualization, writing—review and editing. S.I.A.: methodology, software, validation, writing—review and editing. A.G.U.: conceptualization, validation, formal analysis. D.U.O.: conceptualization, methodology, validation, formal analysis.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to H.U. or D.U.O.

Reprints and permissions information is available at www.nature.com/reprints.

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