Combinatorial Long-Short Term Memory– Group Interaction in Contribution Method for Estimation of Lethal Dose of Pesticides

Lebna Djari[1], Hayat Zerrouki[2], Saber Kouadri[3], KhadraMokadem[4], Zineb Ghiaba[5], Mohamed Lakhdar Belfar[6]

[1] Kasdi Merbah Ouargla University, VPRS Laboratory, B.P. 511, 30000, Ouargla, Algeria

[2] Kasdi Merbah Ouargla University, VPRS Laboratory, B.P. 511, 30000, Ouargla, Algeria

[3] Kasdi Merbah Ouargla University,Laboratory of Water and Environment Engineering in Sahara Milieu (GEEMS), Ouargla, Algeria

[4] Kasdi Merbah Ouargla University, VPRS Laboratory, B.P. 511, 30000, Ouargla, Algeria

[5] Kasdi Merbah Ouargla University, VPRS Laboratory, B.P. 511, 30000, Ouargla, Algeria

[6] Kasdi Merbah Ouargla University, VPRS Laboratory, B.P. 511, 30000, Ouargla, Algeria, mbelfar@gmail.com

Received: 12/2022 **Published: 12/2022**

Abstract:

Humans are continuously exposed to a variety of chemicals , including pesticides , many of which are potentially toxic and have carcinogenic effects . Determining the human toxicity of chemicals remained a challenge due to the large resources required to evaluate a chemical in vivo. In this study , a hybrid model that combined the long-short term memory (LSTM) and Group Interaction Contribution (GIC) was used to predict the LD_{50} toxicity in rats with chemical pesticides , which was developed using a database relatively large consisting of 303 pesticides belonging to different chemical groups . The architecture of LSTM-MLR hybrid model was carefully selected by testing different number of hidden neurons and different number of training iterations in order to avoid the model over fitting . The parameters selection process revealed that the model should be developed using 300 hidden neurons and trained with 1000 iterations . The developed model with best selected parameters achieved a very interesting results and a high accuracy in the prediction of LD_{50} in testing phases with 0.8888 , 0.1312 , 0.3622 , 0.2926 , 20.64 , and 39.99, for R^2 , MSE , RMSE , MAE , RAE, and RRSE , respectively.

Keywords: Long-Short time Memory (LSTM), Group –Interaction Contribution (GIC), lethal dose (LD50), Pesticides, Prediction

Tob Regul Sci. ™ 2022;8(1): 3412-3425
DOI: doi.org/10.18001/TRS.8.1.257

1. Introduction

Recent years have witnessed a noticeable increase in the population (Samsidar et al .2018; Zibaee et al . 2020) and with the terrible spread of pests and their invasion of agricultural crops (Zibaee et al . 2020). The need to provide food has increased (Nayak et al .2020) which prompted people to use all means and methods to confront these pests (Sànchez et al .2011). Pesticides are the best and fast easy way to get rid of them. They are defined as a substance or group of chemicals that are used to control or eliminate harmful organisms (Alavanja et al .2013; Dahiri et al . 2021) .

Pesticides are widely used around the world (Ye et al .2013; Shanta et al . 2018) with different types and quantities (Wu et al .2011) to increase yield and improve food quality (Ramadan et al . 2016) , as well as to combat the spread of vector-borne diseases such as malaria and dengue fever (Pirhadi et al . 2021; Kim et al . 2017). However , its excessive use and indiscriminate dealing with it (Wang et al .2017) led to an increase in pests both quantitatively and qualitatively (Larramendy et al .2014) . In addition to the emergence of several environmental problems , manifested in the pollution of various ecosystems (Ayanda et al .2014), the presence of pesticides and their residues in the ecosystem have also been associated with a wide range of risks that may affected humans and their surroundings(Oluwole et al .2009; Narenderan et al . 2020) , as they are biologically active toxins (Wu et al .2011) where their toxicity depended on the mode of action of the active ingredients. They can also affected living organisms depended on dose and duration of exposure.

The World Health Organization (who) estimated that between one and five million cases of pesticide 20,000 deaths of poisoning were recorded (Kishi et al . 1995) . This number of poisoning cases calls for increased toxicity tests to evaluate any of these chemicals safe for the environment and humans (Gadaleta et al . 2019) . In the early stages of toxicology testing , pesticide risks were usually identified and assessed based on the LD for acute oral toxicity in rodents , which was a standard piece of information required to classify chemicals in terms of the potential risk posed to human health after acute exposure (Rositsa et al .2010). The lethal dose (LD_{50}) was known to kill 50% of experimental animals within 24 hours of exposure and the test was done through the skin or orally(Alberga et al .2019) , as the traditional experimental tests of toxicity of a large number of chemicals were expensive in terms of time and money and sacrifice a number of chemicals and animals (Polash et al .2019) .Recently, it has been observed that researchers tend to use machine learning methods to predict the toxicity of chemical compounds. For example, prediction of the toxicity of silico chemicals has become popular as a faster and inexpensive alternative, eliminating the need of more animal experiments (Limbu et al . 2021) .

In this regard , several computational toxicology models have been developed as viable approaches to reduce both the cost and the number of animals used in the evaluation of experimental toxicity . Several studies have been published dealing with several models of systemic toxicity to rodents, where organophosphates have received special attention. This study

highlighted the usefulness of descriptors such as lipophilic , and the roots of the width , as their results showed that the ANN model was better than the MLR model depending on the correlation factor where the R^2 of the ANN model was equal to 0.95 and the correlation factor of the MLR was equal to 0.86.

2. Experimental

2.1. Data Set:

The database compounds were sorted and 303 compounds were selected from the family of insecticides . The validity of the values in the database was also verified by reviewing several scientific manuscripts that confirmed the same values of the lethal dose of the same compound. Table 1 showed the repartition of the database adopted in this study , while figure 1 showed the general shape of the data base . The data set was randomly divided into a training set represented 80% and a test set represented 20% of the total database.

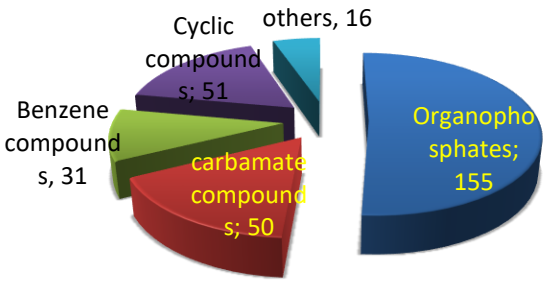

Fig.1 The ratios of compounds to the nervous system of insects

2.2. Group interaction contribution method

In the present study, molecular structure was related to the lethal dose (LD_{50}) of pesticides through a two-level estimation: first-order contribution and second-order contribution. This was carried out according to concepts established in previous works owed to Marrero and Pardillo (Marrero et al . 1999) as well as Constantinou and Gani (Constantinou et al . 1995) . Pesticides were divided into groups based on how their interactions influenced the lethal dose (LD50), which was the property of interest .The model took into account the sub-structures (groups) , their frequency of occurrences , and their interactions for each of the pesticides studied . The number of groups represented in an pesticides was defined as: 0, when the group did not appear in the IL molecule; and *n*, when the group appeared *n* times.

2.3. Long-Short time memory (LSTM) model

The Long-Short time memory or LSTM model appeared for the first time in 1997 by Sepp Hochreiter and Jürgen Schmidhuber in order to solve the problem of store information over extended time interval with recurrent back propagation (Hochreiter et al .1997). This model has the ability to merge and save the impact of model steps during the modeling process at long and

short range. The original LSTM model formed with multiple cells called C, cell for every step in the data set. This cell contained the information needed of long term memory. Figure 1 presented the architecture of LSTM cell, where x presented the input and y presented the output of the cell.

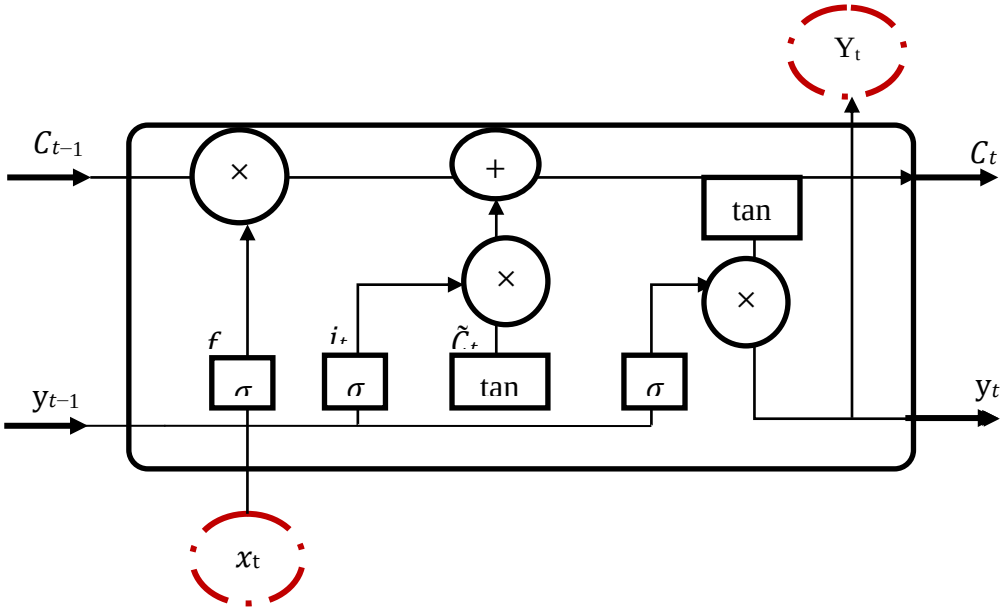

Fig.2. Basic LSTM layer structure for the time step t-1 to t+1, with a detailed calculation illustration shown in the LSTM cell at time step (Xiang et al . 2020).

The process in each LSTM cell consisted of the updates of six parameters, the equations used in the update process were presented in equations 1 – 6.

$$f_t = (Wf \cdot [y_{t-1}, x_t] + b_f) \quad (1)$$

$$i_t = \sigma\,(W_i \cdot [y_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh\,(W_C \cdot [y_{t-1}, x_t] + b_C) \quad (3)$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (4)$$

$$o_t = \sigma\,(W_o \cdot [y_{t-1}, x_t] + b_o) \quad (5)$$

$$y_t = o_t \times \tanh\,(C_t) \quad (6)$$

Each cell start it process by calculating the forget gate (f_t), this parameter produce information about the amount of previous cell state should be forgotten. The calculation at this stage applied on the current input (x_t) and the results obtained from the previous cell (y_{t-1}) using sigmoid and linear equations. The weights (W) and biases (b) have different values from a cell to another. After that, the input gate with the second parameter should be calculated. This parameter specified which information need to be remembered and added it to the current cell stat. here, a sigmoid function with a linear relation on x_t and $y_{t-1}$was used. The next step was the use of tanh

function with a linear relation to calculate the candidate of new cell state ($\tilde{C}t$) based on x_t and y_{t-1}. Consequently, the cell state C_t was updated. As a preparation step for the final output, the sigmoid function with a linear relation was involved again to calculate the output parameter (o_t) using x_t and y_{t-1}. Finally, to get the output final result of the current cell (y_t), a multiplication of o_t and the tanh function value of cell state C_t were generated.

In this work , the problem treated was a regression nature, where the use of the original LSTM model was not possible. For that, and in order to benefit from the advantages of LSTM model in the treatment of regression problems, both LSTM and linear regression were merged through fully connected layer. This operation was made by MATLAB environment, where the Deep Network Designer application was used. Four types of layers were involved to create our model; first, a "sequence input layer" was selected. After that, the "LSTM layer" was posed to receive the inputs. At third position, a "fully connected layer" was used as a link between LSTM layer and the fourth layer which was a "linear regression layer".

Figure 2 presented a screen shot from the Deep Network Designer interface, where figure 3 presented the methodology adopted in the training, testing and best parameters selection of the used model.

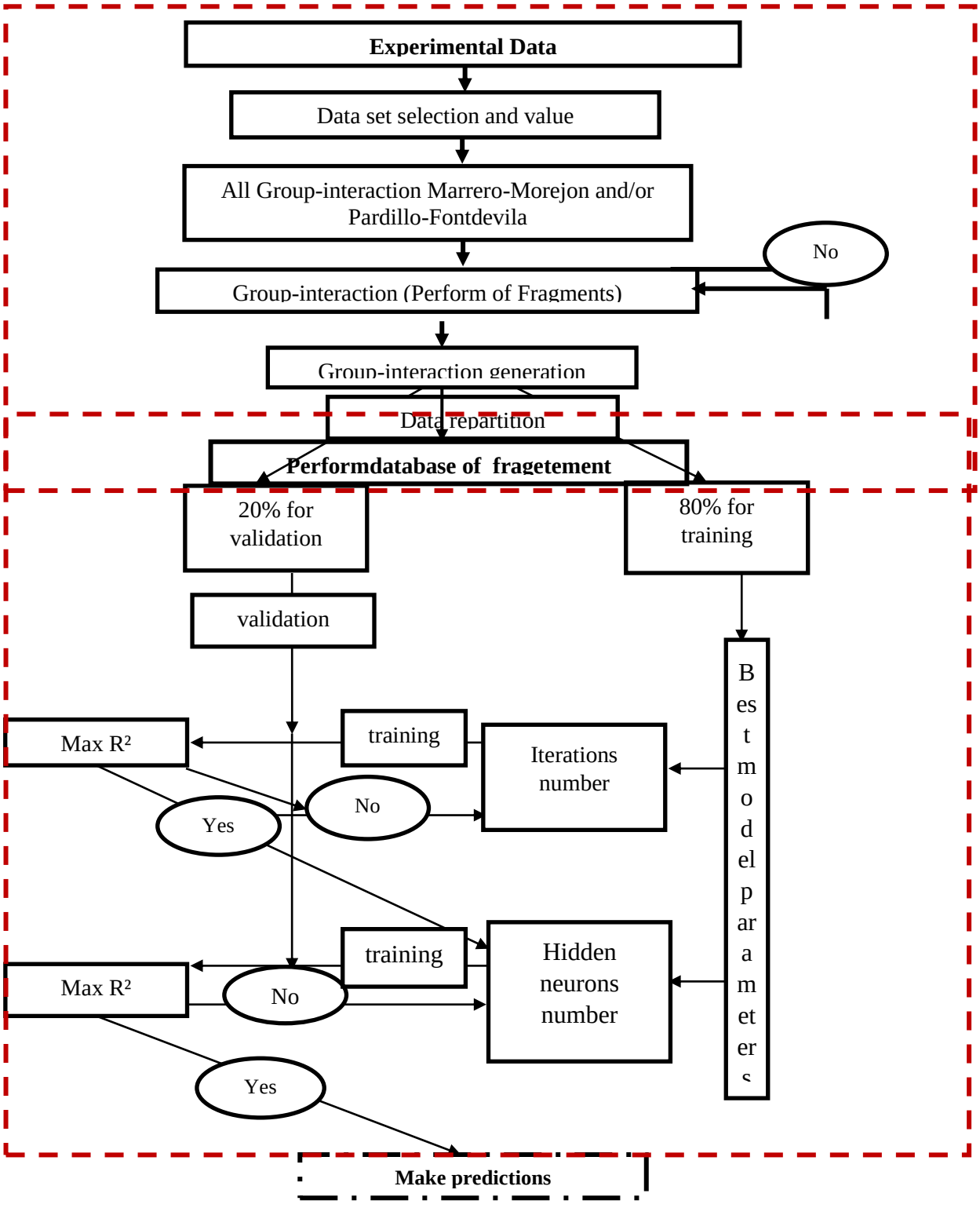

Fig.3. The methodology adopted in training, testing and best parameters selection for the used model

Throughout the analysis, the actual LD_{50} data and the modeled values were compared. The following statistical measures have been chosen to determine the accuracy of the models, the

Root Mean Squared Error (RMSE), mean absolute error (MAE), the relative absolute error (RAE) and the root of relative squared error (RRSE) (Malone et al .2017; Kouadri et al .2021).

1. All settings set as follows:
2. $LD50_A^i$: The measured value of LD_{50}.
3. $LD50_P^i$: Predicted value of LD_{50}.
4. $LD50^-$: The mean value of LD_{50}, and N was the total number of data points.

1. *The Root Mean Squared Error (RMSE)*

The root mean square error between the predicted and actual values was called RMSE. It was given by the following formula:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(LD50_A^i - LD50_P^i)^2} \quad (2)$$

2. *Mean Absolute Error (MAE)*

The mean absolute error evaluated the extent of the errors in a series of predictions without taking their sign into account. It was an estimation of the absolute differences between the expected and observed values on the test sample and defined as follows:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|LD50_P^i - LD50_A^i| \quad (3)$$

3. Relative absolute error (RAE)

The total absolute error was normalized by dividing it by the total absolute error of the basic indicator in the relative absolute error.

$$RAE = \left|\frac{LD50_A^i - LD50_P^i}{LD50_P^i}\right| \times 100 \quad (4)$$

4.The Root of Relative Squared Error (RRSE)

The total squared error was normalized by dividing it by the total squared error of the basic indicator in the relative squared error. The error was reduced to the same dimensions as the predicted quantity by taking the square root of the relative squared error.

$$RRSE = \frac{\sqrt{\sum_{i=1}^{N}(LD50_P^i - LD50_A^i)^2}}{\sqrt{\sum_{i=1}^{N}(LD50_A^i - LD50^-)^2}} \quad (5)$$

3. Results and discussion:

3.1. *Iteration number selection*

This work presented the application of hybridization between LSTM and linear regression models for the aim of estimating the LD50. For that, the parameters set of the model was initialized with 17 different variants. First, the number of hidden neurons left as was in the initial settings which are 200 hidden neurons, where the change set was made at the iterations number. The performance of deferent models was estimated based on six indicators R^2, MSE, RMSE, MAE, RAE, and RRSE, Table 1 presented the performance of the models in both training and testing phases during the adjustment of iteration number. The results shown that the R^2 in the test phase getting close to 1 (ideal value) when the number of iterations grow up, this relation found to be not continuous because after 4000 iterations the value of R^2 getting down which indicated the presence of over fitting of model training. Even if the best R^2 (0.9024) founded to be in the 4000 iterations, the rest of performance indicators values was not optimal at that number of iterations, where we found that the optimal values during test phase for MSE, RMSE, MAE, RAE and RRSE were 0.15151365, 0.38924755, 0.33047161, 22.4283518 and 42.9717669, respectively at the 1000^{th} iteration. Figure 1 presented a radar chart of performance indicators of developed models. The RRSE added as a number and not a percentage to make the graph clearer. Also, the RAE was eliminated in this graph due to the presence of RRSE from one side and due to the marge of values of RAE which was relatively high compared to other indicators in the other side. The radar chart presented five corners; each corner presented a performance indicator where each model presented with a line connected each point from the indicator's axis. Form figure 1, the worst model with 7500 iterations was presented with a red line, on the other hand, the best results founded in the model of 1000 iterations was presented with the yellow line.

Table 1

Performance of LSTM-MLR model based on deferent iterations number.

N Iterations	Phase	R^2	MSE	RMSE	MAE	RAE	RRSE
100	Training	0.3533	0.54142213	0.73581392	0.59485468	36.8050	80.7426
	Test	0.6228	0.3713544	0.60938855	0.46668745	25.4478169	67.2746
200	Training	0.9283	0.06965091	0.26391459	0.17101692	13.8597262	28.960
	Test	0.8644	0.24612074	0.49610558	0.39563399	26.606952	54.7685
500	Training	0.9927	0.00618226	0.07862734	0.04911338	4.19833612	8.62797
	Test	0.8531	0.1596028	0.39950319	0.33591673	22.8928005	44.1039
1000	Training	0.9969	0.00271887	0.05214279	0.02798223	2.87860711	5.7217
	Test	0.8531	**0.1515136 5**	**0.3892475 5**	**0.3304716 1**	**22.428351 8**	**42.9717**

2000	Training	0.9986	0.00118337	0.03440022	0.01801298	2.26346643	3.7748
	Test	0.8295	0.18863881	0.43432569	0.38924586	25.516273	47.9482
4000	Training	0.9996	0.00029449	0.01716079	0.00960214	1.20520909	1.8830
	Test	0.902 4	0.58463087	0.76461158	0.69258479	37.4508009	84.4108
5000	Training	0.9998	0.00013092	0.01144212	0.00699853	0.79174575	1.2555
	Test	0.8884	0.58771179	0.76662363	0.6943039	37.9923067	84.6329
7500	Training	1	6.5061E-05	0.00806605	0.00677901	0.47038928	0.8851
	Test	0.8499	1.79957091	1.34148087	1.15652729	56.3790423	148.095 4
10000	Training	1	0.00030013	0.01732412	0.01639324	0.91923389	1.9010
	Test	0.81	0.28362206	0.53256179	0.47064405	29.422167	58.7932

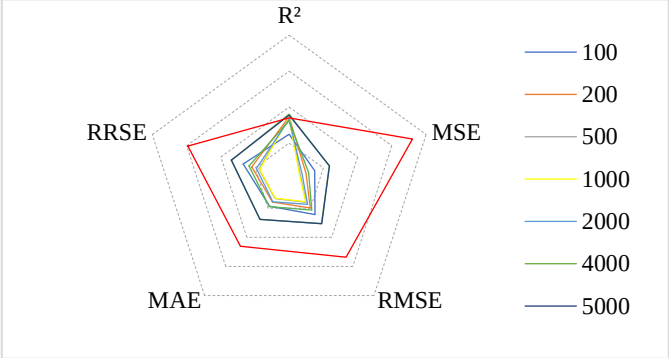

Fig4. LSTM performance based on deferent number of iterations

Hidden neurons selection

After the selection of the best number of iterations and fixed it in 1000, the same procedure made for the selection of the best number of neurons in the hidden layer was performed. Table 2 presented the results of eight models with different architectures, where the number of neurons in the hidden layer varied from 50 to 400 with a step of 50 during training and test phases. The results revealed that the architecture has an impact on the performance of LSTM model. A positive relationship was registered between the number of neurons in the hidden layer and the fitting of the model during the training phase. In the test phase, the augmentation of number hidden neurons from 50 to 150 and from 350 to 400 found to be negatively influenced the model's performance. Otherwise, the variation of the same parameter from 200 to 300 found to be helping the performance of models where the best results achieved with 300 neurons in the hidden neurons with performance indicators equal to 0.1312294, 0.362256, 0.2926325, 20.645653, and 39.991981 for MSE, RMSE, MAE, RAE, and RRSE, respectively.

Figure 5 depicted radar chart with different models based on their architecture and the corresponding performance indicators. From the figure 5 it's clear that the worst model was the

model with 150 neurons in the hidden layer, which was presented with the black line, where the best model was presented with the red line which was the model of 300 neurons in the hidden layer.

By comparing the best models from the first step (table1) and the second step (step2), an improvement in performance of model was recorded after selecting the best number of iterations and the best number of neurons in the hidden layer.

Table 2

Performance of LSTM-MLR model based on deferent Hidden neurons number.

N Hidden neurons	phase	R^2	MSE	RMSE	MAE	RAE	RRSE
50	train	0.9984	0.0013891	0.0372707	0.0177668	1.6592803	4.0898
	test	0.7598	0.2059261	0.4537908	0.370817	21.235603	50.0971
100	train	0.9994	0.0005378	0.0231898	0.012788	1.439076	2.5446
	test	0.8996	0.637451	0.7984053	0.7053828	36.923771	88.1415
150	train	0.9995	0.0003874	0.0196834	0.010759	1.3733947	2.1599
	test	0.8992	1.1687901	1.081106	0.9412616	46.31738	119.3508
200	train	0.9996	0.0002945	0.0171608	0.0096021	1.2052091	1.8831
	test	**0.9024**	0.5846309	0.7646116	0.6925848	37.450801	84.4108
250	train	0.9997	0.0002478	0.0157423	0.0090348	1.1253405	1.7274
	test	0.8606	0.2297078	0.4792784	0.4260715	27.675192	52.9109
300	train	0.9997	0.0002177	0.0147543	0.0084126	1.0634029	1.61902
	test	**0.8888**	**0.1312294**	**0.362256**	**0.2926325**	**20.645653**	**39.9919**
350	train	0.9998	0.000204	0.0142838	0.008195	1.0128264	1.567401
	test	0.8621	0.1369794	0.3701073	0.3194732	21.995087	40.8587
400	train	0.9998	0.0002043	0.0142934	0.0084362	1.0029066	1.5684
	test	0.8499	0.1450867	0.3809025	0.3310468	22.317919	42.0505

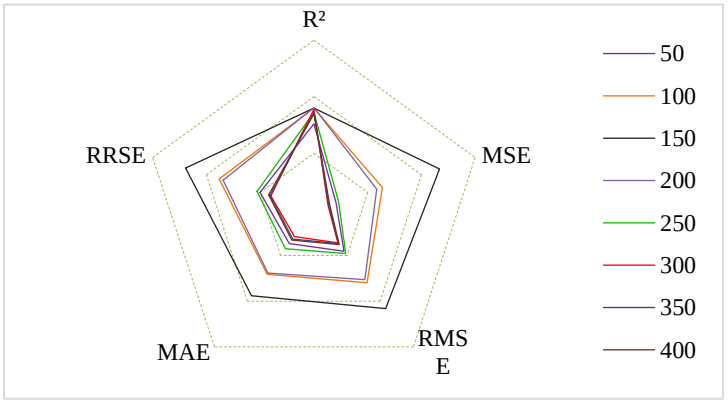

Fig.5. LSTM performance based on deferent number of hidden neurons

The best model composed from 300 neurons in the hidden layer and trained with 1000 iterations, the time series of actual values of LD50 with green circles and the predicted values of LSTM model with orange
circles was presented in figure 6 part "a", also the part "b" of figure 6 presented a scatter plot present the points between actual values and predicted values with green circles, where the red line represent the trend line of projected pointes, where the R^2 founded to be equal to 0.8888.
The results showed that this new model represents an excellent alternative to the classical techniques and an accurate and reliable model for assessing the toxicity of chemical compounds, where the statistical parameters were obtained, in testing phases with 0.8888, 0.1312, 0.3622, 0.2926, 20.64 and 39.99for R^2, MSE, RMSE, MAE, RAE and RRSE, respectively.

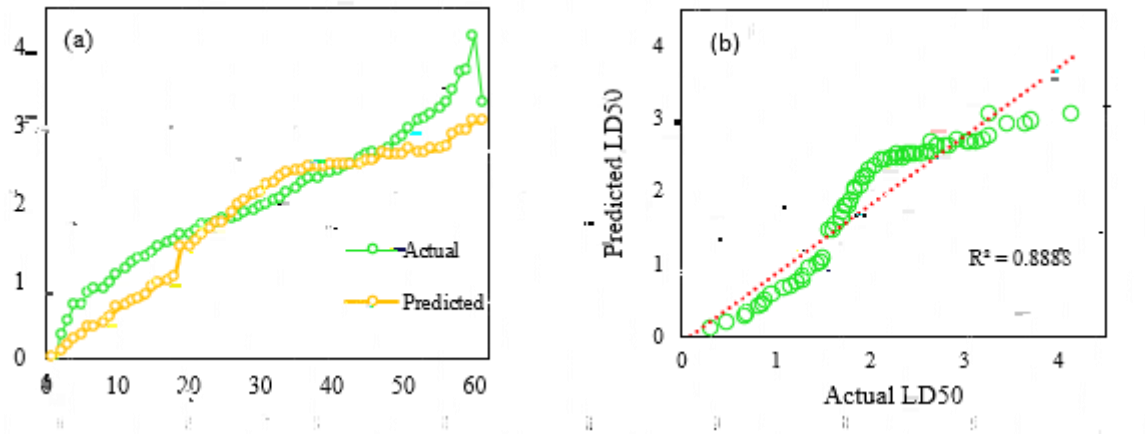

Fig.6. Time series plot (a), and scatter plot (b) of LSTM-MLR model with optimal parameters

The comparison of these finding with those of previous studies mentioned that the modelling of LD50 by using LSTM model gives the best results with optimal values and improvement in performance. Similarly, Martin et al predicted the value of LD50 using a model of linear discrimination analysis LDA and hierarchical clustering model HC, where the prediction accuracy was R^2 = 0.47 and R^2 = 0.50, respectively (Martin et al .2017), while zhu et al used a QSAR model and achieved a prediction accuracy of R^2=0.42 (Zhu et al .2009). Another study found a prediction accuracy of R^2 = 0.86 (Fikri et al .2019) .

Conclusion

The new model used in this study combined a long-short term memory model (LSTM) with contribution of interaction to predict the values of oral LD_{50} of rats based on a database of 303 chemical compounds of insecticides with neurotoxicity. The best model included 300 neurons in the hidden layer and 1000 iterations of training. Based on the results, the architecture has an influence on the performance of LSTM model. The LSTM yielded the greatest results with optimal values and

improved performance as a result of simulation.

References

[1] Alavanja, M. C., Ross, M. K., & Bonner, M. R. (2013). Increased cancer burden among pesticide applicators and others due to pesticide exposure *A: A* Cancer Journal for Clinicians, 63(2), 120-142

[2] Alberga, D., Trisciuzzi, D., Mansouri, K., Mangiatordi, G. F., & Nicolotti, O. (2019). Prediction of acute oral systemic toxicity using a multi finger printsimilarityapproach. Toxicological Sciences, 167(2), 484-495.

[3] Ayanda, O. I., Olasehinde, G. I., & Ajayi, A. A. (2014). Toxicity of sublethal concentrations of glyphosate and paraquat herbicide in the Africancatfish (Clariasgariepinus). International Journal of Agriculture & Biology, 6(20), 1359-1364.

[4] Constantinou L., & Gani R. , (1994).New Group Contribution Method for Estimating Properties of Pure Compounds". AIChE Journal, 40,10,, 1697-1710.

[5] Dahiri, B., Martín-Reina, J., Carbonero-Aguilar, P., Aguilera-Velázquez, J. R., Bautista, J., & Moreno, I. (2021). Impact of Pesticide Exposureamong Rural and UrbanFemale Population. An Overview. International Journal of EnvironmentalResearch and Public Health, *18*(18), 9907. https://doi.org/10.3390/ijerph18189907

[6] Fikri, H., Fechtali, T., & Mamoumi, M. (2019). Evaluation of the model prediction toxicity (ld50) for series of 42 organophosphorus pesticides. Journal of Engineering Studies and Research, 25(1), 30-35.

[7] Gadaleta, D., Vuković, K., Toma, C., Lavado, G. J., Karmaus, A. L., Mansouri, K., ... & Roncaglioni, A. (2019). SAR and QSAR modeling of a large collection of LD 50 rat acute oral toxicity data. Journal of cheminformatics, 11(1), 1-16.

[8] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation,9(8), 1735–1780.

[9] Kim, K. H., Kabir, E., &Jahan, S. A. (2017). Exposure to pesticides and the associated human health effects. Science of the total environment, 575, 525-535.

[10] Kishi, M., Hirschhorn, N., Djajadisastra, M., Satterlee, L. N., Strowman, S., & Dilts, R. (1995). Relationship of pesticide spraying to signs and symptoms in Indonesian farmers. ScandinavianJournal of Work, environment & health, 124-133.

[11] Kouadri, S., Elbeltagi, A., Islam, A. R. M., & Kateb, S. (2021). Performance of machine learning methods in predicting water quality index based on irregular data set: application on Illizi region (Algerian southeast). Applied Water Science, 11(12), 1-20.

[12] Larramendy, M. L., & Soloneski, S. (2014). Pesticides-Toxic Aspects. Intechopen http://dx.doi.org/10.5772/56979

[13] Limbu, S., Zakka, C., & Dakshanamurthy, S. (2021). Predicting Environmental Chemical Toxicity using a New Hybrid Deep Machine Learning Method. ChemRxiv.

[14] Malone, B. P., Styc, Q., Minasny, B., & McBratney, A. B. (2017). Digital soil mapping of soil carbon at the farm scale A spatial downscaling approach in consideration of measured and

uncertain data. Geoderma, 290, 91-99.

[15] Marrero-Morej, J., & Pardillo-Fontdevila, E., (1999). Estimation of pure compound properties using group-interaction contributions. Institute of Chemical Engineers Journal, 45, 615-621.

[16] Martin, T. M., Lilavois, C. R., & Barron, M. G. (2017). Prediction of pesticide acute toxicity using two-dimensional chemical descriptors and target species classification. SAR and QSAR in Environmental Research, 28(6), 525-539.

[17] Narenderan, S. T., Meyyanathan, S. N., & Babu, B. (2020). Review of pesticide residueanalysis in fruits and vegetables. Pre-treatment, extraction and detection techniques. Food Research International, 133, 109141.

[18] Nayak, P., & Solanki, H. (2021). Pesticides and indian agriculture-a review. International Journal of Research-GRANTHAALAYAH, 9(5), 250-263

[19] Oluwole, O., & Cheke, R. A. (2009). Health and environmental impacts of pesticide use practices: a case study of farmers in Ekiti State, Nigeria. International journal of agricultural sustainability, *7*(3), 153-163.

[20] Pirhadi, M., Zeinaly, T., Sadighara, P., & Manouchehri, A. (2021). Residues of Pesticide in Food and Environmental Risk Factors. An Overview of Food Exposure and Its Impact on Human Health. Turkish Journal of Physiotherapy and Rehabilitation, *32*, 3.

[21] Polash, A. H., Takumi, N., & Takeda, S. (2019). Systematic approaches to build predictive models for rat oral toxicity. CICSJ Bulletin, 37(1), 12.

[22] Ramadan, G., Al Jabir, M., Alabdulmalik, N., & Mohammed, A. (2016). Validation of a method for the determination of 120 pesticide residues in apples and cucumbers by LC-MS/MS. Drug testing and analysis, 8(5-6), 498-510.

[23] Rositsa, S., Mojca, F.G., & Andrew, W.L., (2010). Review of QSAR Models and Software Tools for Predicting Genotoxicity and Carcinogenicity. JRC Scientific and Technical Reports, 1-58.

[24] Samsidar A., Siddiquee, S., &Shaarani, S. M. (2018). A review of extraction, analytical and advanced methods for determination of pesticides in environment and foodstuffs. Trends in Food Science &Technology, 71, 188-201

[25] https://doi.org/10.1016/j.tifs.2017.11.011

[26] Sánchez-Bayo, F. (2011). Impacts of agricultural pesticides on terrestrialecosystems. Ecological impacts of toxicchemicals, 63-87,2011.

[27] Shanta, K.. &Hansraj, S. (2018). The Impact of Pesticides on Farmer's Health: A Case Study of Fruit Bowl of Himachal Pradesh. International Journal of Science and Research (IJSR), 3(11),144-148.

[28] Wang, Y., Yang, G., Dai, D., Xu, Z., Cai, L., Wang, Q., &Yu, Y. (2017). Individual and mixture effects of five agricultural pesticides on zebrafish (Daniorerio) larvae. Environmental Science and Pollution Research, 24(5), 4528-4536.

[29] Wu, G., Bao, X., Zhao, S., Wu, J., Han, A., &Ye, Q. (2011). Analysis of multi-pesticide

residues in the foods of animal origin by GC–MS coupled with accelerated solvent extraction and gel permeation chromatography cleanup. Food Chemistry, 126(2), 646-654 doi:10.1016/j.foodchem.2010.10.105

[30] Xiang, Z., Yan, J., & Demir, I. (2020). A rainfall-runoff model with LSTM-based sequence-to-sequence learning. Water resources research, 56(1), e2019WR025326.

[31] Ye, M., Beach, J., Martin, J. W., & Senthilselvan, A. (2013). Occupational pesticide exposures and respiratory health.

[32] International Journal of EnvironmentalResearch and Public Health, 10(12), 6442-6471.

[33] Zhu, H., Martin, T. M., Ye, L., Sedykh, A., Young, D. M., & Tropsha, A. (2009). Quantitative structure– activity relationship modeling of rat acute toxicity by oral exposure. Chemical research in toxicology, 22(12), 1913-1921.

[34] Zibaee, A., & Malagoli, D. (2020). The potential immune alterations in insectpests and pollinatorsafter insecticide exposure in agro ecosystem. Invertebrate Survival Journal, 99-107.