INTRODUCTION

The mitogen-activated protein (MAP) kinases are the key signaling molecules. The MAP kinases contain threonine and serine protein kinase (Wang, Wu, Ai, & Wang, 2015; Wang et al., 1998). They participate in diverse cellular events and are potential targets for affecting inflammation, cancer, and other diseases. The MAP kinase p38 is responsive to environmental stresses and is involved in the production of cytokines during inflammation (Huang et al., 2012; Wang et al., 1998). Numerous diseases including inflammatory and autoimmune diseases such as rheumatoid arthritis (RA) or inflammatory bowel disease (IBD), neurodegenerative conditions, cardiovascular events, and even cancer could be associated with atypical regulation of protein kinase-mediated cell signaling. The progress of chronic inflammation is driven by an amplified systemic occurrence of several proinflammatory cytokines such as tumor necrosis factor-R (TNFR), whose biosynthesis and release is elevated by the activity of mentioned protein kinases, for example, the signal molecule p38 mitogen-activated protein kinase (p38, p38 MAPKa) (Chen et al., 2001; Noble, Endicott, & Johnson, 2004; Schieven, ...
The discovery of specific inhibitors of the p38 block that affect on inflammatory cytokines has generated significant interest in the MAP kinases as drug targets (Lee et al., 1994). Several small molecules p38 MAPK inhibitors have been shown to effectively block the production of IL-1, TNF, and other cytokines in vitro and in animal tests (Adams, Badger, Kumar, & Lee, 2001; Gallagher et al., 1997; Liverton et al., 1999; Wagner & Laufer, 2006).

The resultant toxic effects of chemical compounds on the body exhibit certain specificity which depends on the compound’s chemical structure and reactivity (Nel, Xia, Mäddler, & Li, 2006). Molecular descriptors and chemometrics are a powerful combined tool for studying pharmaceutical, toxicological, and environmental problems. The concept of molecular structure is one the most important concepts in the development of the scientific knowledge. Molecular structures are internal factors governing the physicochemical properties, environmental behavior, and toxicology of organic compounds. Compounds with similar molecular structures should have similar physicochemical properties, environmental fate, and toxicological effects; that is, there are inherent relations between molecular structures and their physicochemical properties, and environmental behavioral and toxicological parameters (Tunkel, Mayo, Austin, Hickerson, & Howard, 2005). The relations can be characterized as mathematical models, termed as structure–activity relationships (SARs) or quantitative structure–activity relationships (QSARs), collectively referred to as (Q)SARs. We have reviewed these concepts in details in other papers (Abdolmaleki & Ghasemi, 2017; Abdolmaleki, Ghasemi, & Ghasemi, 2017a; Abdolmaleki, Ghasemi, & Ghasemi, 2017b; Ahmadi, Khazaei, & Abdolmaleki, 2014). As a matter of fact, the reasoning based on the molecular structure has been the main engine for the great development of molecular sciences. Therefore, molecular descriptors are now playing a key role in scientific research and are applied in modeling several different properties in fields such as analytical chemistry, physical chemistry, medicinal and pharmaceutical chemistry, and environmental and toxicological studies (Chen, Li, Yu, Wang, & Qiao, 2008).

The field of artificial intelligence (AI) or computing intelligence techniques (CI) such as artificial neural networks and fuzzy logic (FL) are the method of choice for structure–activity and structure–property correlation. World Scientific Publishing Co., Inc. The AI-based methods are spanning a wide variety of applications of theoretical modeling and simulation. In recent years, it has been applied successfully in many cases such as knowledge discovery and data mining as well as statistical analysis and machine learning. These techniques received considerable attention because of its abilities in dealing with imprecision and uncertainty and in learning to achieve tractability, robustness, and low solution cost. They mainly consist of fuzzy logic (FL), artificial neural networks, probabilistic reasoning, and evolutionary optimization (Ciaramella et al., 2005; Hou, Su, & Chang, 2008; Katritzky et al., 2001, 2010; Lin & Lee, 1996). Neural networks are often used in conjunction with optimization techniques for feature selection, ranging from simple greedy approaches such as forward selection or backward elimination, and other AI techniques such as adaptive neuro-fuzzy inference systems (ANFIS), ant colony, and particles swarm (Agrafiotis, Cedeno, & Lobanov, 2002; Buyukbingol, Sisman, Akylidiz, Alparslan, & Adejare, 2007; Loukas, 2001; Mwense et al., 2006). FL has been applied in different fields, such as computational biology, physics, automated control, and decision-making support. The underlying idea in FL applications is that people are capable of making decision using imprecise or uncertain knowledge, whereas traditional computer algorithms require precise information. An important advantage of using fuzzy models is that they are capable of incorporating knowledge from human experts naturally and conveniently, while traditional models fail to do so. Other important properties of fuzzy models are their ability to handle nonlinearity and interpretability feature of the models. FL is the most popular constituent of the CI area since fuzzy systems are able to represent human expertise in the form of IF antecedent THEN consequent statements. In this domain, the system behavior is modeled through the use of linguistic descriptions (Babuška & Verbruggen, 2003; Efe & Kaynak, 2001; Nelles, 2001; Paripour, Ferrara, & Salimi, 2017; Zadeh, 1996; Zhou & Gan, 2008).

The integration of these methodologies that exploit the strength of each collection and synergistically is a driving force to synthesize intelligent systems. Therefore, in the present work, a hybrid method consisted of ANFIS, MLR, and ANN is developed for diverse set of an important group of medicinal compounds. In the following sections, the theory of the proposed methods is briefly described. Then, the results of the proposed system and the comparison results are illustrated. Finally, concluding remarks are made.

## 2 | METHODOLOGICAL BACKGROUND

Normally, in a QSAR/QSPR study, an optimization algorithm is used to select the descriptors containing the most information about the given property or activity. CI methods can be complementary to previous approaches and can be used to search efficiently over many disciplines of biomedicine and biochemistry.

### 2.1 | Artificial neural networks (ANNs)

Artificial neural network is an information-processing paradigm that inspired by biological nervous systems. ANN is the type of artificial intelligence that attempts to imitate the way a human brain works and learn the relationships between certain inputs and outputs from experience. An ANN
consists of several layers of a large of highly interconnected computational units called neurons working in union to solve specific problems. They are considered as information processing systems that have the abilities to learn, recall, and generalize from training data. Basically, the artificial neural network consists of neurons, simple processing elements, which are activated as soon as their inputs exceed certain thresholds. The neurons are arranged in layers which are connected so that the signals at the input are propagated through the network to the output (Gath & Geva, 1989).

ANNs are useful tools in QSAR/QSPR studies and particularly in cases where it is difficult to specify an exact mathematical model for describing a given structure-property relationship. In fact, ANN is a popular strategy for nonlinear modeling in QSAR studies and some researchers have indicated its capability for drug discovery (Deeb & Clare, 2007; Deeb & Drabb, 2010; Deeb & Goodarzi, 2010; Deeb & Hemmateenejad, 2007; Ghasemi, Mehridehnavi, Pérez-Garrido, & Pérez-Sánchez, 2018). Most of these works used neural networks based on the backpropagation learning algorithm, which has some disadvantages such as local minimum, slow convergence, time-consuming nonlinear iterative optimization, and difficulty in explicit optimum network configuration. In contrast, the parameters of radial basis function neural networks (RBFNNs) can be adjusted by fast linear methods. It has advantages of short training times and is guaranteed to reach the global minimum of error surface during training. The optimization of its topology and learning parameters are easy to be implemented (Walczak & Massart, 2000; Yao et al., 2004). RBFNNs can be described as a three-layer feed-forward structure and consist of three layers: an input layer, hidden layer, and an output layer. The input layer does not process the information; it only distributes the input vectors to the hidden layer. The transfer function in the hidden layer of RBF networks is called the kernel or basis function. For a detailed description, the reader is referred to references. Each hidden layer unit represents a single radial basis function, with associated center position and width. Each neuron on the hidden layer employs a radial basis function as a nonlinear transfer function to operate on the input data (Yao et al., 2004).

2.2 Adaptive neuro-fuzzy inference systems (ANFIS)

Adaptive neuro-fuzzy inference systems are synthesized by an appropriately integrating the neural and fuzzy system interpretations. Therefore, the resulting hybrid combination inherits the numeric power of NN as well as the verbal power of FL (Efe, Kaynak, & Wilamowski, 2000; Jung, Sun, & Mizutani, 1997; Li, Huang, & Chen, 2004). An ANFIS structure having m inputs and the single output with product inference rule and first-order Sugeno model can be described as in (Noble et al., 2004) with $f_i$ being described as in the rule consequent. The structural view of such a system is illustrated in Figure 1. The rule structure for an ANFIS utilizing first-order Sugeno model has the following representation: IF $u_i$ is $U_i$ AND $u_2$ is $U_2$ AND … AND $u_m$ is $U_m$ THEN $f_i = \sum_{i=1}^{m} \mu_i u_i + \sum_{i=m+1}^{n} \mu_i u_i \cdot \tau$. When the consequent part of the rule structure is compared with that of rules in SFS architecture, it is seen that the polynomial representation of the decision introduces higher parametric flexibility extending the realization capability. ANFIS structure has been utilized with gradient-based training strategies for identification of nonlinear systems (Efe & Kaynak, 1999). An in-depth discussion is given with numerous examples on the use of ANFIS structure (Jang et al., 1997).

3 EXPERIMENTAL DATA AND DATA ANALYSIS METHODS

3.1 Datasets and descriptor generation

The dataset consisted of an experimental dataset for pyridinylimidazole-based compounds that were taken from the

FIGURE 1 Structure of an ANFIS
study Laufer et al. as shown in Table 1 (Laufer, Hauser, Domeyer, Kinkel, & Liedtke, 2008). The biological data (IC$_{50}$) based on research of Laufer et al. used in this study were expressed in μM. In this case, the log-transformed values did not improve significantly the model’s prediction ability. At first, the geometry optimizations of all 3D structures headed to energy minima that were accomplished by using AM1 method of hyperchem Program. Many descriptors calculated to study the similarity or dissimilarity of compounds and used to predict the relationships between physicochemical properties and biological activities by chemoffice (CS ChemOffice 2005 molecular modeling software version 9) and MM+ (Molecular Modeling Pro plus (MMPP) version 6.0 (ChemSW Inc.)) programs. Total descriptors were utilized to remove the intercorrelating descriptors (redundancy) and reduce the multicollinearity (Buyukbingol et al., 2007).

In the ANFIS feature selection step, the dataset was partitioned into a training set (odd-indexed samples) and a checking set (even-indexed samples) to select the set of inputs that most influence the IC$_{50}$ and overfitting problem was monitored by plotting resulting data. In the modeling step both MLR and ANN, 75% data are selected for the training set and 25% for the check set. Individual elements are selected randomly. The data values (each data) were extracted and written in separate training and testing data files. The MATLAB (version 7.1, MathWorks, Inc) and Neural Networks Toolbox were used for neural networks practicability.

### 3.2 ANFIS feature selection

A critical issue with many biological datasets is the overwhelming number of possible features that could be used as input to a model. Standard statistical approaches can be used to reduce the number of features, but in many cases, thousands of features may still remain. And it may also be the case that the researcher is most interested in identifying the key nonlinear relationships between features that are useful in an output decision, and reducing the feature space using linear regression methods may not be well suited. In such cases, the researcher can use the evolutionary process to evolve the selection of features to be used as input to the model concurrent with the optimization of the model itself. Such a method rapidly identifies useful collections of features while simultaneously producing an optimized model. Therefore, an exhaustive search was used within the available inputs to determine the set of the most influential input attribute in predicting the activity of mentioned drug or output. When the training and checking errors are comparable, this implies that there is no overfitting. By this way, an ANFIS model was constructed for each combination and trains it and reports the performance achieved.

We have obtained 31 descriptors out of 49 descriptors from the datasets which were utilized to find the significant descriptors. The selected descriptors (after removing input attributes with RMSE > 0.6) for ANFIS model with one variable as shown in Figure 2. We have used the constrain of RMSE > 0.6 for avoiding overfitting and overtraining. We can simply select models with two, three, or four variables directly since they have the least errors as shown in the plot. However, this will not necessarily be the optimal combination of two or three inputs that result in the minimal training error. To verify this, we repeated the search for the optimal combination of two, three, and four input attributes.

### 3.3 Development of the models

Building good predictive models is difficult, and the results are not always satisfactory. However, the performance of a model may be acceptable in some domain, whereas insufficient in another. Recent studies show that combined models minimize uncertainty and produce more robust predictors. Here, two approaches of MLR and ANN methods were considered based on selected descriptors by ANFIS method.

#### 3.3.1 MLR method

The stepwise multiple linear regression procedure was used for model generation. In the stepwise multiple linear regression, a linear equation is produced, but all independent variable is not used. Each variable is added to the equation at a time, and a new regression is performed. The new term is retained only if equation passes a test for significance. Three descriptors selected by ANFIS were used as input for the regression analysis using the stepwise procedure. The stepwise method implemented in the software package of SPSS (version 11.5, SPSS Inc.). As a first step, a correlation matrix was performed for these three descriptors calculated for each molecule. Inspection of this matrix did not show a considerable correlation ($R \geq 0.90$) between them. Of course, this dataset remained after eliminating of one outlier (4) from dataset; outliers $\geq 2SD$ where $SD$ is standard deviation measured by SPSS. Among the more suitable model obtained, the best MLR model was chosen for further evaluation. The best MLR model consists of two descriptors. The two parameters appearing in this model were the Homo and LogP. The main goals of generating the MLR model were developing a calibration model for the prediction of IC$_{50}$ as a linear model.

#### 3.3.2 ANN method

ANNs are useful tools in QSAR/QSPR studies and particularly in cases where it is difficult to specify an exact mathematical model for describing a given structure–property relationship. Most of these works used neural networks
**TABLE 1** Biological activity of different pyridinylimidazole-based compounds include the following: (I) 2-Halogen-substituted pyridin-4-yl- and quinolin-4-ylimidazole derivatives with a 4-fluorophenyl moiety at imidazole C4 or C5. (II) 2-Halogen-substituted pyridin-4-ylimidazole derivatives with a 3-trifluoromethylphenyl moiety at imidazole C4 or C5. (III) Heteroarylalkylaminopyridines (2-thioimidazole derivatives with a 4-fluorophenyl ring). (IV) Heteroarylalkylaminopyridines, (cyclo)alkylaminopyridines and (cyclo)alkylaminopyridines with an additional polar residue (2-thioimidazole derivatives with a 3-trifluoromethylphenyl ring). (V) (Cyclo)alkylaminopyridines and (cyclo)alkylaminopyridines with an additional polar group (2-thioimidazole derivatives with a 4-fluorophenyl ring). (VI) 2-(Thi)oxopyridinylimidazoles. (VII) Tetrasubstituted aminopyridinylimidazoles

| No. | R<sup>1</sup> | R<sup>2</sup> | p38α IC<sub>50</sub> ± SEM (μM) | No. | R<sup>1</sup> | R<sup>2</sup> | p38α IC<sub>50</sub> ± SEM (μM) |
|-----|---------------|---------------|----------------|-----|---------------|---------------|----------------|
| 1   |               | CH<sub>3</sub> | 0.559 ± 0.002 (n = 2) | 2   |               | C<sub>2</sub>H<sub>5</sub> | 0.213 (n = 1) |
| 3<sup>b</sup> |               | CH<sub>3</sub> | 1.04 (n = 1) | 4 |               | | 0.258 (n = 1) |
| 5<sup>b</sup> | | CH<sub>3</sub> | 0.069 ± 0.024 (n = 4) | 6 | | C<sub>2</sub>H<sub>5</sub> | 0.372 ± 0.108 (n = 4) |
| 7 | | CH<sub>3</sub> | 0.411 ± 0.056 (n = 4) | 8<sup>b</sup> | | | 2.90 ± 1.94 (n = 4) |

(Continues)
| No. | R<sup>1</sup> | R<sup>2</sup>   | p38α IC<sub>50 ± SEM</sub> (μM) | No. | R<sup>1</sup> | R<sup>2</sup>   | p38α IC<sub>50 ± SEM</sub> (μM) |
|-----|-------------|-------------|---------------------------------|-----|-------------|-------------|---------------------------------|
|     |             |             | 1.09 ± 0.396 (n = 4)            | 10  | F           | -C<sub>4</sub>H<sub>9</sub> | 1.95 ± 0.345 (n = 4)            |
| 11  | F           | -C<sub>2</sub>H<sub>5</sub> | 0.657 ± 0.139 (n = 4)           |     |             |             |                                 |
|     |             |             |                                 |     |             |             |                                 |
| 12<sup>b</sup> | -C<sub>6</sub>H<sub>5</sub> |             | 0.139 ± 0.001 (n = 2)           | 23  |            | -C<sub>4</sub>H<sub>9</sub> | 0.013 ± 0.002 (n = 6)           |
| 13<sup>b</sup> |            |            | 0.04 ± 0.005 (n = 4)            | 24  |            |             | 0.348 ± 0.054 (n = 4)           |
| 14  |            |            | 0.068 ± 0.014 (n = 6)           | 25  |            |             | 2.15 ± 0.46 (n = 8)             |
| 15  |            |            | 0.028 ± 0.003 (n = 6)           | 26  |            |             | 1.34 ± 0.311 (n = 14)           |
| 16  |            |            | 0.047 ± 0.018 (n = 4)           | 27  |            |            | 0.017 ± 0.007 (n = 4)           |
| 17<sup>b</sup> |            |            | 0.173 ± 0.028 (n = 4)           | 28<sup>b</sup> |            | -N<sub>H</sub> | 0.07 ± 0.012 (n = 4)           |

(Continues)
| No. | $R^1$ | $R^2$ | $p38\alpha$ IC$_{50}$ ± SEM (μM) | No. | $R^1$ | $R^2$ | $p38\alpha$ IC$_{50}$ ± SEM (μM) |
|-----|-------|-------|---------------------------------|-----|-------|-------|---------------------------------|
| 18  | ![Image](image1.png) | 0.082 ± 0.031 (n = 6) | 29$^b$ | ![Image](image2.png) | 0.098 ± 0.024 (n = 4) |
| 19  | ![Image](image3.png) | 0.171 ± 0.083 (n = 4) | 30$^b$ | ![Image](image4.png) | 0.119 ± 0.056 (n = 4) |
| 20  | ![Image](image5.png) | 0.046 ± 0.011 (n = 4) | 31  | ![Image](image6.png) | 0.06 ± 0.02 (n = 4)   |
| 21  | ![Image](image7.png) | 0.023 ± 0.002 (n = 6) | 32  | ![Image](image8.png) | 0.314 ± 0.085 (n = 4) |
| 22  | ![Image](image9.png) | 0.017 ± 0.007 (n = 6) |     |     |        | (Continues)                     |
| 33$^b$ | ![Image](image10.png) | 0.282 ± 0.108 (n = 4) | 38  | ![Image](image11.png) | 0.239 ± 0.028 (n = 4) |
| 34$^b$ | ![Image](image12.png) | 0.199 ± 0.089 (n = 4) | 39  | ![Image](image13.png) | 0.191 ± 0.028 (n = 4) |
| 35  | ![Image](image14.png) | 0.248 ± 0.110 (n = 4) | 40  | ![Image](image15.png) | 0.036 ± 0.004 (n = 4) |

(Continues)
| No. | R<sup>1</sup> | R<sup>2</sup> | p38α IC<sub>50</sub> ± SEM (μM) | No. | R<sup>1</sup> | R<sup>2</sup> | p38α IC<sub>50</sub> ± SEM (μM) |
|-----|-------------|-------------|-------------------|-----|-------------|-------------|-------------------|
| 36  | H<sub>3</sub>C | | 0.093 ± 0.02 (<i>n</i> = 4) | 41 | CH<sub>3</sub> | | 0.034 ± 0.003 (<i>n</i> = 8) |
| 37  | H<sub>3</sub>C | | 0.049 ± 0.014 (<i>n</i> = 4) | 42 | | | 0.061 ± 0.015 (<i>n</i> = 4) |
| 43  | | | 0.011 ± 0.0009 (<i>n</i> = 10) | 52 | | |  |
| 44<sup>b</sup> | | | 0.08 ± 0.018 (<i>n</i> = 4) | 53 | | | 0.126 ± 0.041 (<i>n</i> = 4) |
| 45  | | | 0.009 ± 0.001 (<i>n</i> = 8) | 54 | | | 0.011 ± 0.002 (<i>n</i> = 10) |
| 46  | | | 0.153 ± 0.017 (<i>n</i> = 6) | 55 | | | 0.023 ± 0.004 (<i>n</i> = 6) |
| 47  | | | 0.02 ± 0.004 (<i>n</i> = 6) | 56 | | | 0.026 ± 0.003 (<i>n</i> = 6) |
| 48  | | | 0.013 ± 0.005 (<i>n</i> = 4) | 57 | | | 0.02 ± 0.004 (<i>n</i> = 6) |
| 49<sup>b</sup> | | | 0.031 ± 0.006 (<i>n</i> = 8) | | | | 0.199 ± 0.089 (<i>n</i> = 4) |

<sup>b</sup> Compound (V)
| No. | R<sup>1</sup> | R<sup>2</sup> | p38α IC<sub>50</sub> ± SEM (μM) | No. | R<sup>1</sup> | R<sup>2</sup> | p38α IC<sub>50</sub> ± SEM (μM) |
|-----|--------------|-------------|-------------------------------|-----|--------------|-------------|-------------------------------|
| 50<sup>b</sup> |  |  | 0.02 ± 0.005 (n = 4) | 58<sup>b</sup> |  |  | 0.045 ± 0.004 (n = 4) |
| 51<sup>b</sup> |  |  | 0.09 ± 0.036 (n = 10) | 59 |  |  | 0.015 ± 0.002 (n = 4) |
| 60 |  |  | 0.035 ± 0.005 (n = 4) | 64 |  |  | 0.515 ± 0.019 (n = 4) |
| 61 |  |  | 0.006 ± 0.002 (n = 4) | 65 |  |  | 0.005 ± 0.0003 (n = 4) |
| 62 |  |  | 0.01 ± 0.001 (n = 4) | 66<sup>b</sup> |  |  | 0.003 ± 0.0005 (n = 4) |
| 63 |  |  | 0.034 ± 0.006 (n = 6) |  |  |  |  |
| 67 | O |  | 0.436 ± 0.086 (n = 4) | 72 | O |  | n. t. (impure) |
| 68 | O |  | 0.639 ± 0.083 (n = 4) | 73 | O |  | 0.641 ± 0.044 (n = 4) |
| 69 | O |  | 0.033 ± 0.01 (n = 4) | 74 | S |  | n. t. (impure) |
| No. | R¹ | R² | p38α IC₅₀ ± SEM (μM) | No. | R¹ | R² | p38α IC₅₀ ± SEM (μM) |
|-----|----|----|----------------------|-----|----|----|----------------------|
| 70  | O  |    | 0.424 ± 0.095        | 75  | S  |    | 0.087 ± 0.026        |
|     |    |    | (n = 4)              |     |    |    | (n = 4)              |
| 71  | O  |    | 0.542 ± 0.145        |     |    |    |                      |
|     |    |    | (n = 4)              |     |    |    |                      |
| 76  |    | CH₃| 0.96 ± 0.42          | 84  |    | CH₃| 0.042 ± 0.0005       |
|     |    |    | (n = 4)              |     |    |    | (n = 4)              |
| 77b |    | CH₃| 0.07 ± 0.03          | 85  |    | CH₃| 0.235 ± 0.051        |
|     |    |    | (n = 4)              |     |    |    | (n = 4)              |
| 78b |    | CH₃| 0.61 ± 0.02          | 86b |    |     | 0.038 ± 0.005        |
|     |    |    | (n = 4)              |     |    |    | (n = 6)              |
| 79  |    | CH₃| 0.04 ± 0.01          | 87  |    | CH₃| 0.003 ± 0.0002       |
|     |    |    | (n = 4)              |     |    |    | (n = 4)              |
| 80  |    | CH₃| 0.051 ± 0.004        | 88  |    | CH₃| 0.211 ± 0.017        |
|     |    |    | (n = 4)              |     |    |    | (n = 4)              |
| 81  |    | CH₃| 0.359 ± 0.097        | 89  |    | CH₃| n.t. (impure)        |
|     |    |    | (n = 8)              |     |    |    |                      |
| 82  |    | CH₃| 0.292 ± 0.023        | 90  |    | CH₃| 0.146 ± 0.002        |
|     |    |    | (n = 4)              |     |    |    | (n = 4)              |
| 83  |    | CH₃| 0.485 ± 0.025        |     |    |    |                      |
|     |    |    | (n = 4)              |     |    |    |                      |

Note. n.t.: not tested.

* n: number of experiments. b Selected as test set.
based on the backpropagation learning algorithm, which has some disadvantages such as local minimum, slow convergence, time-consuming nonlinear iterative optimization, and difficulty in explicit optimum network configuration (Walczak & Massart, 2000; Yao et al., 2004). In this work, the input parameters \( \text{Diam} \), \( \text{Homo} \), and \( \text{LogP} \) were normalized to simplify of descriptors effect on IC\(_{50}\) within the range \((-1, 1)\). The feed-forward network was formed with six neurons in the hidden layer and one neuron in the output layer. \( \text{Radbas} \) and \( \text{tansig} \) transfer functions were chosen in the hidden layer and output layer, respectively. The learning algorithm is Leverberg–Marquardt with mean squared error as performance function. This algorithm is more efficient than the basic backpropagation (BP) algorithm and is highly recommended as a first-choice supervised algorithm, although it does require more memory than other algorithms.

4 | RESULTS AND DISCUSSION

The main goals of the present work were as follows: (a) to build regression models based on ANFIS system feature selection, to reveal the relationship of novel 2,4,5- and 1,2,4,5-substituted 2-thioimidazoles, (b) to achieve a better understanding of the physicochemical basis of inhibitor activity and its mechanism in these drug, and (c) to compare the ability of the linear (MLR) and nonlinear (ANN), chemometrics techniques in predicting the activity behavior of a diverse set of 2-thioimidazoles derivatives. To fulfill these goals, one needs a very diverse dataset. As can be seen from Table 1, a dataset consisting of 87 drug candidates with very different structures was chosen to develop the appropriate models.

4.1 | ANFIS results

ANFIS to evolve the selection of features rapidly identifies useful collections of features while simultaneously producing an optimized model. The goal of the ANFIS is to find a model or mapping that will correctly associate the inputs (structural descriptors) with the output (biological activity). By this procedure, the most significant descriptors (inputs) in the fuzzy system were automatically identified. The ANFIS is simulated using the MATLAB Fuzzy Logic Toolbox.

ANFIS turned out to be less appropriate for biological applications, since the number of adjustable network parameters grows exponentially with the number of molecular descriptors, owing to the network's complex architecture. Therefore, for practical use, the number of descriptors was restricted to at most 12 (removing input attributes with RMSE > 0.6) inputs. The plot and results from the exhaustive search (Figure 2) clearly indicate that the input attribute \( \text{LogP} \) is the most influential. We can select more than one input attribute to build the ANFIS model. We examined ANFIS models with two, three, and four variables. The plot demonstrates the result of selecting three inputs, in which \( \text{Diam} \), \( \text{Homo} \), and \( \text{LogP} \) were selected as the best combination of three input variables. However, the minimal training (and checking) error does not reduce significantly from that of the best four-input model, which indicates that the newly added attribute does not improve the prediction much. For better generalization, we always prefer a model with a simple structure. ANFIS dynamically constructs the initial (input) and final (output) membership functions (MFs) based on the nature of data. The degree an object belongs to a fuzzy set is denoted by a membership value between 0 and 1. The \text{genfis1} function generates an initial FIS from the training data, which is then fine-tuned by ANFIS to generate the final model.
Figure 3 indicates the minimal checking error occurs at about epoch 30, which is indicated by a circle. ANFIS was trained for 28 epochs with gauss2mf as a proper MF. Therefore, we put three-input ANFIS for further exploration. Notice that the checking error curve goes up after 30 epochs, indicating that further training overfits the data and produces worse generalization.

The ANFIS prediction can be compared against a linear regression model by comparing their respective RMSE values against checking data. Thus, the performance of the ANFIS model by different MFs including of gauss2mf, gaussmf, and trimf was checked with a linear regression model. RMSE against checking data for ANFIS models was 0.535, 0.646, and 0.573, respectively. The higher value of RMSE (6.547) in the linear regression and the near-zero value of the RMSE of the ANFIS method which indicates the appropriate descriptors shows the superiority of the ANFIS method than linear regression. The final selected descriptors in our system belong to three groups: steric, electronic, and thermodynamic descriptors. As noted before, the correlation matrix for the highest significant three descriptors showed no significant intercorrelated descriptors.

The attractive features of an ANFIS include easy to implement, fast and accurate learning, strong generalization abilities, excellent explanation facilities through fuzzy rules, and easy to incorporate both linguistic and numeric knowledge for problem-solving (Roy & Chakraborty, 2013).

4.2 | MLR and ANN analysis

4.2.1 | MLR results

Linear regression methods such as MLR and PLS have been widely used for obtaining predictive and descriptive QSAR/QSPR models in our research group (Ahmadi et al., 2014; Ghasemi, Abdolmaleki, Asadpour, & Shiri, 2008; Ghasemi, Asadpour, & Abdolmaleki, 2007; Ghasemi & Saaidpour, 2007; Ghasemi, Saaidpour, & Brown, 2007; Rouhollahi, Shafieyan, & Ghasemi, 2007). In this QSAR study, the stepwise MLR analysis was employed on the training dataset to establish the quantitative regression model. MLR method provided a useful equation linking the structural features to the IC50 of the compounds as the following form:

$$IC_{50} = -13.2064 (\pm 2.2733) + -1.5920 (\pm 0.2827) \text{Homo} + 0.1261 (\pm 0.0474) \text{LogP}$$

This QSAR model can be used to predict the activity of compounds based on these parameters. In the proposed model, one of two variables (Homo or highest occupied molecular orbital energy) is related to the electronic properties of the molecules (e.g., the partial charge distribution or the...
electronegativities of atoms) and another descriptor, Log P, is related to the thermodynamic structure of molecules.

Validating QSAR with external data (i.e., data not used in the model development) is the best method of validation. The predictive power of the regression model developed on the selected training set is estimated on the predicted values of prediction set chemicals. Multiple regression analysis of these data yielded a pretty good fit ($R = 0.8050$) for chemicals activity. The residuals values as shown in Figure 4 obtained by the MLR modeling as differences of experimental and predicted IC$_{50}$ values. The model obtained is good, bearing in mind the great variety of functionality: amino, alkyl, hydroxyl, carbonyl, and carboxyl groups, and aromatic rings. Furthermore, existence of a trend in the residual plot as has shown in Figure 4, for this data, is maybe a reason for nonlinearity of this dataset. It is well-known that biological systems are inherently nonlinear and dynamic, and also, cumulative probability plot (Figure 5) shows a nonlinearity in the distribution of dataset; then, this plot verified the fact of nonlinearity of dataset.

### 4.2.2 | ANN results

Inspection of results in prior section shows this dataset is a typical nonlinear regression problem. This problem can be tackled by using the neural learning algorithm. Since there might be a strongly nonlinear component in the relationship between IC$_{50}$ and the three descriptors (resulted by ANFIS), QSAR modeling was also performed using ANNs to get the ANN counterpart to Equation (1). The learning capability of the network was tested with varying the transfer functions in hidden and output layer and the number of neurons in the hidden layer as shown in Table 2. It observed that network generalizes well for six neurons in the hidden layer and radbas and tan sig transfer functions as the optimal value. The mean squared error is found to be 0.0197 for training set and 0.0231 for testing set. The training was terminated after 323 epochs.

Satisfactory results obtained with the Levenberg–Marquardt method using RBFNN. In case of radial basis network, the hidden layer consists of radbas transfer function (Duch & Jankowski, 2001). The conjugate method; Levenberg–Marquardt method is extremely fast. The advantage of the Levenberg–Marquardt is that it converges faster around the minimum and gives more accurate results. Its only drawback is that it requires more memory than the backpropagation with momentum method. In contrast, backpropagation NNs, the parameters of RBFNNs, can be adjusted by fast linear methods. It has advantages of short training times and is guaranteed to reach the global minimum of error surface during training. The optimization of its topology and learning parameters are easy to be implemented (Walczak & Massart, 2000; Yao et al., 2004).

### 4.3 | Comparison and measures of performance

This section presents some results obtained through the performing and training of several neural network architectures with the gradient descent backpropagation algorithm, using real data as network input features. If there is enough data available, it is preferable to divide it into a training set and an external test set. The test set is used for validating the model after training and will function as a completely new unseen dataset. The model performance on the external test set acts as a generalization measure. In the cases when there was no external test set supplied in the original data, the data were sorted according to the y-values and every value was assigned to the test set randomly. To be able to determine the accuracy of the results and compare the predictive abilities of the methods on the dataset, some kind of performance measure is needed. Two commonly used measures in multivariate

| No. of neuron | Transfer function (hidden layer) | Transfer function (output layer) | Network training function | $R$ (train) | MSE (train) | MSE (test) | No. of epoch |
|---------------|----------------------------------|----------------------------------|---------------------------|------------|------------|------------|--------------|
| 7             | Tansig                            | Tansig                           | Train lm                  | 0.85       | 0.03       | 0.03       | 100          |
| 7             | Radbas                            | Tansig                           | Train lm                  | 0.92       | 0.02       | 0.07       | 79           |
| 8             | Radbas                            | Tansig                           | Train lm                  | 0.87       | 0.03       | 0.02       | 100          |
| 9             | Radbas                            | Tansig                           | Train lm                  | 0.87       | 0.03       | 0.02       | 100          |
| 6             | Radbas                            | Purelin                          | Train lm                  | 0.91       | 0.02       | 0.04       | 85           |
| 6*            | Radbas                            | Tansig                           | Train lm                  | 0.92       | 0.01       | 0.02       | 323          |
| 6             | Tansig                            | Tansig                           | Train lm                  | 0.92       | 0.01       | 0.03       | 239          |
| 7             | Radbas                            | Tansig                           | Train lm                  | 0.93       | 0.02       | 0.04       | 100          |
| 7             | Radbas                            | Tansig                           | Train lm                  | 0.96       | 0.009      | 0.18       | 1000         |
| 7             | Radbas                            | Purelin                          | Train lm                  | 0.97       | 0.001      | 0.23       | 1000         |

*Selected as optimum procedure.
analysis are the mean-square error (MSE) and the determination coefficient or the squared correlation coefficient also referred to as the $R^2$-value.

A summary of the results of prediction performances of proposed methods for train and test data is given in Table 3. The comparison between two methods shows that the ANN method is more effective. Therefore, the lower the MSE values, the neural network has better prediction performance. Similarly, the higher the correlation coefficient ($R$) value, the higher the prediction performances of the model. According to correlation coefficient ($R$) in both trained and tested data, ANN is better predictor model. The correlation coefficient of $\geq 0.90$ for the prediction set showed a good predictive ability for the generated model by RBFNN. As a matter of fact, the prediction of the test set based on selected descriptors by ANFIS revealed its suitability in feature reduction. The results show that the test result is relatively poor for MLR model, as can be seen in its $R$ value. The ANN results obtained in this QSAR study are better than the corresponding MLR results. Apparently, the results were significantly improved in these ANN runs over the MLR analyses, indicating the strong non-linear dependence of inhibition data on the input descriptors used. The strength of the alternative multivariate methods compared to the traditional multilinear regression (MLR) and PLS is their ability to handle complex non-linear data. Comparison of results for MLR and NN show that neural networks can be effectively used in the activity development of pyridinylimidazole-based compounds and gives more satisfactory results for practical problems. In summary, NN was able to predict the test data.

4.3.1 Model validation by external test criteria

Model validation is a critical phase of each QSAR modeling. Commonly, much of the effort has focused on new validation methods for producing superior answers. It has the potential to significantly enhance answers. In computational manner, it could be interpreted in many ways. Roy and Tropsha have introduced interesting statistical tools for validation of acquired QSAR models. Golbraikh and Tropsha through analysis of several datasets showed the assumption of leave-one-out cross-validated $R^2$ more than 0.5 or (LOO $q^2 > 0.5$) is necessary for assessing the predictive ability of model but it is not sufficient. They emphasized that the test set or external validation is the best tactic to form a reliable QSAR model (Golbraikh & Tropsha, 2002). Generally, the following parameters are the most popular criteria for more validation of QSAR/QSPR models:

$R^2 > 0.6$ (2)

$$\frac{(R^2 - R^2_0)}{R^2} < 0.1 \text{ or } \frac{(R^2 - R^2_0)}{R^2} < 0.1$$ (3)

$$0.85 \leq k' \leq 1.15 \text{ or } 0.85 \leq k \leq 1.15$$ (4)

$R^2_0$ and $R^2$ are correlation coefficients for regression through the origin for predicted versus observed and observed versus predicted activities, respectively, and corresponding slopes of regression lines through the origin are $k$ and $k'$. Besides, Roy (Pratim Roy, Paul, Mitra, & Roy, 2009; Tropsha, Gramatica, & Gombar, 2003) proposed another validation statistical criteria, named modified $R^2$ ($R^2_m$) which measure the predictability of a model and is determined as follows:

$$R^2_m = R^2 \left(1 - \sqrt{|R^2 - R^2_0|}\right)$$ (5)

$R^2_m$ value more than 0.5 ($R^2_m > 0.5$) approves that the model possess good external prediction ability.

In this respect, using mentioned statistical tools for the validation of MLR and ANN models, we listed values in Table 4. The results for all criteria in this table suggest high ability of the ANN model in prediction of the bioactivity of candidate drugs.

4.4 Explanation of descriptors

The best algorithms of QSARs modeling are those that are simple, transparent, easily interpretable, and easily portable. A transparent model can be defined as one that is based on

| Validation tools | MLR | ANN |
|------------------|-----|-----|
| $R^2 \geq 0.6$   | 0.6428 | 0.9134 |
| $\frac{(R^2 - R^2)}{R^2} < 0.1$ | 0.2 | 0.05 |
| $0.85 \leq k' \leq 1.15$ | 1.91 | 1.04 |
| $R^2_m = R^2 \left(1 - \sqrt{|R^2 - R^2_0|}\right)$ | 0.46 | 0.71 |
fundamental physicochemical properties with a clear and unambiguous statement of how the model has been formulated (Cronin & Schultz, 2003). Such a model is capable of mechanistic interpretation. The transparent characteristics are usually achieved by proper mathematical algorithms (Schultz & Cronin, 2003). In other words, a QSAR model should be interpretable in terms of the parameters employed. Interdisciplinary integration with subjects such as biochemistry and toxicology can deepen the understanding of the modes of toxicity actions and then improve mechanistic interpretabilities of QSAR models (Cronin & Schultz, 2003; Schultz & Cronin, 2003).

The essential interactions of pyridinylimidazole inhibitors with the ATP-binding cleft are briefly summarized as follows: (a) hydrogen donor/acceptor functions of the 2-aminopyridyl residues within the hinge region (mainly gaining activity); (b) space-filling lipophilic aryl residues binding to the hydrophobic back pocket (also hydrophobic region, mainly gaining selectivity); (c) interactions with the hydrophobic front region (also hydrophobic region, gaining both activity and selectivity); (d) further interactions with both the sugar pocket and the phosphate binding region (importance less clear, preferred positions to modify physicochemical properties) (Tong et al., 1997; Wang et al., 1998; Wilson et al., 1997).

As noted above, the process parameters that have significant effects on the inhibition activity of these drugs include Diam, Homo, and LogP descriptors. Diam attribute of a compound depends mainly on the molecular size and symmetry, and it influences on intermolecular interactions. This physical property for a crystal is governed by the hydrogen bonding ability of the molecules, the molecular packing in crystals (effects from molecular shape, size, and symmetry), and other intermolecular interactions such as charge transfer and dipole–dipole interactions in the solid phase (Ghasemi et al., 2008). Therefore, it influences on solubility and solubility controls $IC_{50}$. The polarization of a molecule by an external electric field seems to correlate with shape, size. It is possible; make inductive interactions in the molecule. Thus, the most significant property of the molecular polarizability is the relation to the molecular bulk or molar volume. Polarizability values have been shown to be related to hydrophobicity and thus affect on biological activities. Furthermore, the electronic polarizability of molecules shares common features with the electrophilic superdelocalizability (Karelson, Lobanov, & Katritzky, 1996). This suggests that, with larger substituents, the interactions between receptor residues and compounds initially are enhanced, resulting in an increase in activity. However, as substituents become too large, unfavorable steric repulsions between ligand and receptor begin to dominate, resulting in a decrease in activity (Chiu & So, 2004).

The earlier studies established small (methyl) substituents located at the exocyclic sulfur atom significantly improved inhibition of activity in both the isolated kinase assay and cytokine release from the human whole blood. This greater inhibitory potential was attributed to better entry of the inhibitor molecule into the binding cleft of p38, leading to stronger enzyme–drug interactions (Laufer, Striegel, & Wagner, 2002). For example, in the kinase activity assay, the ethylsulfanyl substituent of compound 6 proved to be nearly bioisosteric to the well-defined methylsulfanyl substituent (e.g., at compound 1, divergence factor of 1.7), additionally confirming the idea of stronger binding of the 4-fluorophenyl/pyridin-4-yl pharmacophore inside the ATP site, when substituents at the exocyclic sulfur are not bulky. For compound 8, the large benzylsulfanyl moiety at the 2-position of the imidazole core diminished the p38 inhibitory potency in contrast to its sterically less demanding ethylsulfanyl analog 6 (3.8-fold) and particularly with respect to the “small” methylsulfanyl derivatives ML3375, 1, and 3. Consequently, the least displaced orientation of those modified “bulker” drug molecules within the active site was responsible for the consistently reduced activity of the 3-trifluoromethylphenyl-substituted imidazoles (Laufer et al., 2008).

Our modeling shows that the Homo energy is an important factor in QSAR models. Energies of the Homo and Lumo are very popular quantum chemical descriptors. It has been shown that these orbitals play a major role in governing many chemical reactions (i.e., formation of charge transfer complexes) and determining electronic band gaps (as stability index) in solids. A large Homo–Lumo gap implies high stability for the molecule in the sense of its lower reactivity in chemical reactions. The Homo–Lumo gap has also been used as an approximation to the lowest excitation energy of the molecule. The energy of the Homo is directly related to the ionization potential and characterizes the susceptibility of the molecule toward attack by electrophiles. The energy of the Lumo is directly related to the electron affinity and characterizes the susceptibility of the molecule toward attack by nucleophiles. The concept of hard and soft nucleophiles and electrophiles has been also directly related to the relative energy of the Homo/Lumo orbitals. Hard nucleophiles have a low-energy Homo; soft nucleophiles have a high-energy Homo; hard electrophiles have a high-energy Lumo and soft electrophiles have a low-energy Lumo (Karelson et al., 1996). According to the frontier molecular orbital theory (FMO) of chemical reactivity, the formation of a transition state is due to an interaction between the frontier orbitals (Homo and Lumo) of reacting species. Frontier orbital electron densities on atoms provide a useful means for the detailed characterization of donor–acceptor interactions. Thus, the majority of chemical reactions takes place at the position and in the orientation where the overlap of the Homo and Lumo of the respective reactants can reach a maximum. In the case of a donor molecule, the Homo density is critical to the charge transfer (electrophilic electron density), and in the case of an...
acceptor molecule, the Lumo density is important (nucleophilic electron density). These indices have been employed in this QSAR study to describe drug–receptor interaction sites. According to the polyelectronic perturbation theory of Klopman and Hudson, drug–receptor interactions are under either charge or orbital control. Thus, the net atomic charges characterize electrostatic interactions, while the donor superdelocalizability characterizes the covalent component of the interaction (Karelson et al., 1996).

The dataset (Table 1) shows the introduction and wide variation of substituents at the Pyridine part of these 2-thioimidazoles bridged by different heteroatoms (NH, O, S). The strength of the H-bonding between the pyridine-N of the inhibitor and the amide-NH of Met109 is influenced by the electron donating or withdrawing capacity of the substituents. The p38 inhibitory potency has been correlated with the electron density at the respective heterocyclic ring nitrogen (Laufer, Wagner, Kotschenreuther, & Albrecht, 2003). An electronegative halogen atom located at the 2-position of the pyridine-4-yl moiety significantly decreased the biological activity. Fluoropyridine 1, for example, exhibited a 2.8-fold worse inhibitory potency toward the isolated enzyme (Laufer et al., 2008).

In agreement with our previous study, the hydrophobicity parameter (LogP) plays an important role in QSAR models (Ghasemi, Abdolmaleki, & Mandoumi, 2009). The most popular quantitative scale to measure the lipophilicity of compounds is the logarithm of the partition coefficient (called the log $P$ parameter) between 1-octanol and water, introduced in detail by Hansch and Leo (Mannhold, Poda, Ostermann, & Tetko, 2009). Partition coefficients have been shown to correlate with measures of biological activity in a very wide variety of experimental systems, ranging from simple protein binding to animal and human in vivo effects. It is a very important indicator of transport and permeation through membranes, interaction with biological receptors and enzymes, toxicity, and biological potency. In environmental sciences, the hydrophobicity is often used to predict solubility, the bioconcentration factor, and the organic adsorption coefficient ($K_{oc}$) (Roberts, 2002). This is presumably because hydrophobic effects are important not only in the intermolecular interactions that occur between a drug and its target site but also in the distribution of a compound within a biosystem and its interaction with competing for binding sites. The results of Livingston’s studies indicated that the dye lipophilicity can be explained by hydrophobic and polarity dye structural parameters (Livingstone, 2000). The positive standardized coefficient for LogP parameter is in accordance with physical considerations—compounds with higher hydrophobicities have stronger interactions with the target site and thus enhance the activity of chemicals.

Hydrogen bonding between the pyridine ring N of the inhibitor and the backbone amide-NH of Met109 in the otherwise lipophilic linker region (Ala51, Thr106, His107, Leu108, and Met109) is essential for the biological activity of the vicinal diaryl-substituted heterocyclic compound class. Appropriate amino substituents at the 2-position of the pyridine can enhance activity toward unsubstituted (e.g., ML3375) and toward halogenated pyridine derivatives, as they are able to form a second H-bond with the linker region next to the existing one (here, between the exocyclic 2-amino group as H-donor and the main chain carbonyl oxygen of Met109 as H-acceptor). On the other hand, an electron donating amino group should increase the electron density at the pyridine, thereby intensifying the affinity of the inhibitor for the enzyme. Moreover, by adequate lipophilic side chains, further interactions with the surface-exposed front region (hydrophobic region) can be achieved. Therefore, in general, sufficiently lipophilic aryalkylamino residues favorably affected biological activity (compounds 13–23, 18). For condensed aromatic ring systems (15, 16, 19), the spatial geometry seemed to play a crucial role; however, multiple separate aromatic cycles in the aliphatic side chain resulted in increased IC$_{50}$ values (24, 25) (Laufer et al., 2008).

A typical three-dimensional (3D) QSAR technique that ultimately allows one to design and predict activities of molecules, for which the foundation is that interaction force fields of series bioactive molecules with the same receptor are similar. Even 3D structures of receptors are unclear, one can deduce the properties of receptors, design new chemicals, and quantitatively estimate activities of chemicals by studying circumambient interaction force fields of bioactive molecules and quantifying bioactivities (Cramer, Patterson, & Bunce, 1988; Marshall & Cramer, 1988).

## 5 | CONCLUSIONS

The 3D structure of drugs to their biological properties has been increased knowledge of the 3D structure of biological molecules. Molecular structural descriptors used in 3D QSARs model capable of mechanistic evaluation. The establishment of QSARs based on the proper analysis and understanding of mechanisms, and vice versa, well-established QSAR models, maybe facilitate mechanism interpretations. In this study, it has been shown that the use of ANFIS offers a feasible method for the optimization of the knowledge base of fuzzy logic controllers. ANFIS process to evolve the selection of features rapidly identifies useful collections of features while simultaneously producing an optimized model, and it is very useful for automated feature reduction and selection tool in QSAR study. We showed that using a real case, the ability and feasibility of the application of hybridization of methods ANFIS _ neural networks a tool further efficiency than ANFIS _ MLR, in biological data. In agreement with the fact,
biological systems are inherently nonlinear and dynamic. It seems these hybrid methods are suitable for facilitating a mechanistic understanding of adsorption, distribution, metabolism, excretion, and toxicity (ADMET) properties and perfectly are a proper interpretation of such systems in QSAR research. These methods can provide reasonable solutions to the researcher in fast time.

ACKNOWLEDGMENTS
Declared none.

CONFLICT OF INTEREST
The authors confirm that this article content has no conflict of interest.

ORCID
Azizeh Abdolmaleki https://orcid.org/0000-0002-1816-2142

REFERENCES
Abdolmaleki, A., & Ghasemi, J. B. (2017). Dual-acting of hybrid compounds-A new dawn in the discovery of multi-target drugs: Lead generation approaches. Current Topics in Medicinal Chemistry, 17, 1096–1114. https://doi.org/10.2174/156802661666616927151144

Abdolmaleki, A., Ghasemi, J. B., & Ghasemi, F. (2017a). Computer aided drug design for multi-target drug design: SAR/QSAR, molecular docking and pharmacophore methods. Current Drug Targets, 18, 556–575. https://doi.org/10.2174/1389450176666160101120822

Abdolmaleki, A., Ghasemi, F., & Ghasemi, J. B. (2017b). Computer-aided drug design to explore cyclodextrin therapeutics and biomedical applications. Chemical Biology & Drug Design, 89, 257–268. https://doi.org/10.1111/cbdd.12825

Adams, J. L., Badger, A. M., Kumar, S., & Lee, J. C. (2001). 1 p38 MAP kinase: Molecular target for the inhibition of pro-inflammatory cytokines. Progress in Medicinal Chemistry, 38, 1–60. https://doi.org/10.1016/S0079-6468(08)70091-2

Agrafiotis, D. K., Cedeno, W., & Lobanov, V. S. (2002). On the use of neural network ensembles in QSAR and QSIPR. Journal of Chemical Information and Computer Sciences, 42, 903–911. https://doi.org/10.1021/ci0203702

Ahmadi, S., Khazaei, M. R., & Abdolmaleki, A. (2014). Quantitative structure–property relationship study on the intercalation of antican cer drugs with ct-DNA. Medicinal Chemistry Research, 23, 1148–1161. https://doi.org/10.1007/s00044-013-0716-z

Babuška, R., & Verbruggen, H. (2003). Neuro-fuzzy methods for nonlinear system identification. Annual Reviews in Control, 27, 73–85. https://doi.org/10.1016/S1367-5788(03)00009-9

Buyukbingol, E., Sisman, A., Akyildiz, M., Alparslan, F. N., & Adejare, A. (2007). Adaptive neuro-fuzzy inference system (ANFIS): A new approach to predictive modeling in QSAR applications: A study of neuro-fuzzy modeling of PCP-based NMDA receptor antagonists. Bioorganic & Medicinal Chemistry, 15, 4265–4282. https://doi.org/10.1016/j.bmc.2007.03.065

Chen, Z., Gibson, T. B., Robinson, F., Silvestro, L., Pearson, G., Xu, B.-E., … Cobb, M. H. (2001). MAP kinases. Chemical Reviews, 101, 2449–2476. https://doi.org/10.1021/cr000241p

Chen, J., Li, X., Yu, H., Wang, Y., & Qiao, X. (2008). Progress and perspectives of quantitative structure-activity relationships used for ecological risk assessment of toxic organic compounds. Science in China Series B: Chemistry, 51, 593–606. https://doi.org/10.1007/s11426-008-0076-6

Chiu, T.-L., & So, S.-S. (2004). Development of neural network QSPr models for Hansch substituent constants. 2. Applications in QSAR studies of HIV-1 reverse transcriptase and dihydrofolate reductase inhibitors. Journal of Chemical Information and Computer Sciences, 44, 154–160. https://doi.org/10.1021/ci030294i

Ciaramella, A., Donalek, C., Staiano, A., Ambrosio, M., Aramo, C., Benvenuti, P., ... Tagliaferri, R. (2005). Applications of neural networks in astronomy and astroparticle physics. Recent research developments in astronomy and astrophysics, 2, 1–32.

Cramer, R. D., Patterson, D. E., & Bunce, J. D. (1988). Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. Journal of the American Chemical Society, 110, 5595–5597. https://doi.org/10.1021/ja00226a005

Cronin, M. T., & Schultz, T. W. (2003). Pitfalls in QSAR. Journal of Molecular Structure: THEOCHEM, 622, 39–51. https://doi.org/10.1016/S0166-1280(02)00616-4

Deeb, O., & Clare, B. W. (2007). QSAR of aromatic substances: Protein tyrosine kinase inhibitory activity of flavonoid analogues. Chemical Biology & Drug Design, 70, 437–449. https://doi.org/10.1111/j.1747-0285.2007.00578.x

Deeb, O., & Drabih, M. (2010). Exploring QSARs of some analgesic compounds by PC-ANN. Chemical Biology & Drug Design, 76, 255–262. https://doi.org/10.1111/j.1747-0285.2010.01004.x

Deeb, O., & Goodarzi, M. (2010). Exploring QSARs for inhibitory activity of non-peptide HIV-1 protease inhibitors by GA-PLS and GA-SVM. Chemical Biology & Drug Design, 75, 506–514. https://doi.org/10.1111/j.1747-0285.2010.00953.x

Deeb, O., & Hemmateenejad, B. (2007). ANN-QSAR model of drug-binding to human serum albumin. Chemical Biology & Drug Design, 70, 19–29. https://doi.org/10.1111/j.1747-0285.2007.00528.x

Duch, W., & Jankowski, N. (2001). Transfer functions: Hidden possibilities for better neural networks. European Symposium on Artificial Neural Networks Bruges (Belgium), 25-27 April D-Facto public, ISBN 2-930307-01-3, pp. 81–94.

Efe, M. O., & Kaynak, O. (1999). A comparative study of neural network structures in identification of nonlinear systems. Mechatronics, 9, 287–300. https://doi.org/10.1016/S0957-4158(98)00047-6

Efe, M. O., & Kaynak, O. (2001). Variable structure systems theory based training strategies for computationally intelligent systems. Denver, CO: IEEE, 1563–1576. https://doi.org/10.1109/IECON.2001.975526

Efe, M. O., Kaynak, O., & Wilamowski, B. M. (2000). Creating a sliding mode in a motion control system by adopting a dynamic defuzzification strategy in an adaptive neuro fuzzy inference system. Nagoya, Japan: IEEE, 894–899. https://doi.org/10.1109/IECON.2000.972241

Gallagher, T. F., Seibel, G. L., Kassis, S., Laydon, J. T., Blumenthal, M. J., Lee, J. C., ... Adams, J. L. (1997). Regulation of stress-induced cytokine production by pyridinylimidazoles; inhibition of CSBP
Paripour, M., Ferrara, M., & Salimi, M. (2017). Approximate solutions by artificial neural network of hybrid fuzzy differential equations. *Advances in Mechanical Engineering*, 9, 1–9.

Pratim Roy, P., Paul, S., Mitra, I., & Roy, K. (2009). On two novel parameters for validation of predictive QSAR models. *Molecules*, 14, 1660–1701. https://doi.org/10.3390/molecules14051660

Roberts, D. W. (2002). Application of octanol/water partition coefficients in surfactant science: A quantitative structure–property relationship for micellization of anionic surfactants. *Langmuir*, 18, 345–352. https://doi.org/10.1021/la0108050

Rouhollahi, A., Shafieyan, H., & Ghasemi, J. B. (2007). A QSPR study on the GC retention times of a series of fatty, dicarboxylic and amino acids by MLR and ANN. *Annali di Chimica: Journal of Analytical, Environmental and Cultural Heritage Chemistry*, 97, 925–933. https://doi.org/10.1002/(ISSN)1612-8877

Roy, S., & Chakraborty, U. (2013). *Introduction to soft computing: Neuro-fuzzy and genetic algorithms*. New Delhi, India: Dorling Kindersley, Pvt. Ltd.

Schieven, G. L. (2005). The biology of p38 kinase: A central role in inflammation. *Current Topics in Medicinal Chemistry*, 5, 921–928. https://doi.org/10.2174/1568026054985902

Schultz, T. W., & Cronin, M. T. (2003). Essential and desirable characteristics of ecotoxicity quantitative structure–activity relationships. *Environmental Toxicology and Chemistry*, 22, 599–607.

Tong, L., Pau, S., White, D. M., Rogers, S., Crane, K. M., Cywin, C. L., … Pargellis, C. A. (1997). A highly specific inhibitor of human p38 MAP kinase binds in the ATP pocket. *Nature Structural and Molecular Biology*, 4, 311. https://doi.org/10.1038/nsb0497-311

Troppsha, A., Gramatica, P., & Gombar, V. K. (2003). The importance of being earnest: Validation is the absolute essential for successful application and interpretation of QSPR models. *QSAR & Combinatorial Science*, 22, 69–77. https://doi.org/10.1002/qsar.1612-0218

Tunkel, J., Mayo, K., Austin, C., Hickerson, A., & Howard, P. (2005). Practical considerations on the use of predictive models for regulatory purposes. *Environmental Science & Technology*, 39, 2188–2199. https://doi.org/10.1021/es049220t

Wagner, G., & Lauper, S. (2006). Small molecular anti-cytokine agents. *Medicinal Research Reviews*, 26, 1–62. https://doi.org/10.1002/ (ISSN)1098-1128

Walczak, B., & Massart, D. (2000). Local modelling with radial basis function networks. *Chemometrics and Intelligent Laboratory Systems*, 50, 179–198. https://doi.org/10.1016/S0169-7439(99)00056-8

Wang, Z., Canagarajah, B. J., Boehm, J. C., Kassissia, S., Cobb, M. H., Young, P. R., … Goldsmith, E. J. (1998). Structural basis of inhibitor selectivity in MAP kinases. *Structure*, 6, 1117–1128. https://doi.org/10.1010/s0969-2126(98)00113-0

Wang, Y., Wu, M., Ai, C., & Wang, Y. (2015). Insight into the structural determinants of imidazole scaffold-based derivatives as TNF-α release inhibitors by in silico explorations. *International Journal of Molecular Sciences*, 16, 20118–20138. https://doi.org/10.3390/ijms160920118

Wilson, K. P., McCaffrey, P. G., Hsiao, K., Pazhanisamy, S., Galullo, V., Bemis, G. W., … Su, M. S. (1997). The structural basis for the specificity of pyridinylimidazole inhibitors of p38 MAP kinase. *Chemistry & Biology*, 4, 423–431. https://doi.org/10.1016/S1074-5521(97)90194-0

Yao, X., Panaye, A., Doucet, J.-P., Zhang, R., Chen, H., Liu, M., … Fan, B. T. (2004). Comparative study of QSAR/QSPR correlations using support vector machines, radial basis function neural networks, and multiple linear regression. *Journal of Chemical Information and Computer Sciences*, 44, 1257–1266. https://doi.org/10.1021/ci049965i

Zadeh, L. A. (1996). *Fuzzy sets: Fuzzy sets*. Binghamton, New York: World Scientific Publishing Co., Inc., 394–432. doi:10.1142/AFSAT

Zhou, S.-M., & Gan, J. Q. (2008). Low-level interpretability and high-level interpretability: A unified view of data-driven interpretable fuzzy system modelling. *Fuzzy Sets and Systems*, 159, 3091–3131. https://doi.org/10.1016/j.fss.2008.05.016

How to cite this article: Abdolmaleki A, Ghasemi JB. Inhibition activity prediction for a dataset of candidates’ drug by combining fuzzy logic with MLR/ANN QSAR models. *Chem Biol Drug Des*. 2019;93:1139–1157. https://doi.org/10.1111/cbdd.13511