Using Neural Networks and Ensemble Techniques based on Decision Trees for Skin Permeability Prediction

Emir Bušatlić[1], Ahmed Osmanović[2], Alma Jakupović[2], Jasna Nuhić[3], Adnan Hodžić[4]

[1] Faculty of Pharmacy, University of Sarajevo
[2] Faculty of Engineering and Information Technologies, Department for Genetics and Bioengineering, International Burch University Sarajevo
[3] Faculty of Engineering and Information Technologies, Department for Electrical Engineering, International Burch University Sarajevo
[4] Faculty of Engineering and Information Technologies, Department for Information Technologies, International Burch University Sarajevo

Abstract. Development of an accurate skin permeability model is becoming increasingly important as skin has been more utilized in recent development of drug delivery methods. This paper presents results of development of Artificial Neural Network (ANN) for prediction of skin permeability. The performance of developed ANN was compared to three regression algorithms used in this paper. The prediction of skin permeability is based on three input parameters: molecular weight, partition coefficient - log(P), and melting temperature for each drug. The dataset of 400 samples was used for prediction of skin permeability. Out of that number, 75% was used for training of ANN, and testing of developed ANN was performed on 100 samples from the dataset. During testing, system correctly predicted 76.7%. This dataset was also used as input to three ensemble techniques based on decision trees: REPTree, Bagging, Random SubSpaceDeveloped. It was shown that Bagging algorithm outperformed developed ANN with 81% while RandomSubspace performed at 73.3%. System can be used in laboratory conditions and can be used in the future for drug discovery.

Keywords: skin permeability, prediction, artificial neural network, regression algorithm, intelligent systems

1 Introduction

The skin is the largest organ of human body, which performs many vital functions like thermoregulation, prevention of excess water loss from the body and protection against external chemical, physical, and biological assailants [1]. Adsorption describes the physicochemical accumulation of substances to other components usually of high molecular weight [2]. Process of adsorption can be separated in three steps: penetration - the entering of substances into a particular layer or organ, permeation - the penetration through one layer into another layer, and resorption - uptake of a substance through the vascular system [3].

Skin absorption is a global term that describes the transport of chemicals from the outer surface of skin both into the skin and into the systemic circulation. There is considerable variability in the measurement of skin permeability. Major differences can be in permeability between spaces. There is a little facts known about variation due to age, although the skin structure does change with age. Sex and ethnic background do not seem to be sources of variation in skin permeability [4]. Transdermal therapeutic system require drugs to penetrate into stratum cornea, and into the systemic circulation in sufficient concentrations for the desire therapeutic effect to occur [2].

First systematic investigation in skin permeability was undertaken by Schwenkenbecher who summarized the knowledge accumulated on the subject at the time [5]. Scheuplain et al [6, 7] and Scheuplain and Blank et al. [8, 9] were first to introduce anatomically based physicochemical models describing percutaneous adsorption. Subsequnetly Flynn et al. [10, 11] proposed the following working model of the skin to assess the permeation of chemicals from their physicochemical properties. In 1988 Flynn and Stewart et al. [12] developed an empirical algorithm for obtaining order of magnitude estimates of the permeability coefficients of drug compound from their octanol/water partition coefficients [13]. Data compiled by Flynn et al. [10] (1990) have been subject to independent statistical analyses by several investigators. All of investigators were trying to find equation to calculate partition coefficient (logP), and form a datasets of their results.

For more than last ten years, Artificial Neural Networks (ANN) have received a great deal of attention and they have being touted as one of the greatest computational tools ever developed. ANNs simulate learning and generalization behavior of the human brain through data modeling and pattern recognition for complex multidimensional problems. Just as brain is composed of numerous neurons, neu-

A. Badnjevic (ed.), *CMBEBIH 2017*,
IFMBE Proceedings 62,
DOI: 10.1007/978-981-10-4166-2_8

ral network has similar structure, artificial neurons. Through data modeling and pattern recognition, these structures simulate learning and human brain behavior [14]. There are many types of neural networks, but the most commonly applied ones are forward propagating network trained by error back-propagation developed by Rumehart et al., as mentioned in [15]. There is significant difference between an ANN model and a statistical model. ANN can generalize the relationship between independent and dependent variables without a specific mathematical function [15]. In pharmaceutical sciences structures like this can be used in drug modeling, pharmacokinetics and pharmacodynamics modeling, dosage design, protein structure and function prediction, interpretation of analytical data and in vivo/in vitro correlations and many more.

Studies by Agatonovic-Kustrin et al.[16] and Cheng et al. [17] have been published on use of ANNs in determining the quantitative structure-permeability relationship of penetration across skin or polydimethylsiloxane membranes as the model of skin permeation. Degim et al. [18] analyzed skin permeability of 40 compounds by ANN and compared its predictability with the multiple linear regression model obtained by Pugh et al. [19] the partial charges of the penetrants, their molecular weight, and their calculated octanol-water partition coefficient log(P)oct/w were used as molecular descriptor.

Popular model for predictive modeling in medicine and pharmacy is tree-based methods for regression. In 2009, Geurts et al. [20] made research on decision tree-based methods and they concluded that this method stand out as non-parametric methods that have the unique feature of combining interpretability, efficiency, and, when used in ensembles of trees, excellent accuracy. In review article, Lemon et al. [21] concluded that regression tree is promising research tool for the identification of at-risk populations in public health research and outreach. In 2012, Hu et al. [22] used decision tree-based learning algorithms to predict analgesic consumption and Patient Controlled Analgesia (PCA) control readjustment based on the first few hours of PCA medications.

Table 1. Studies and number of drug samples taken from each study

Studies	Number of drug samples
Degim et al. [18], Prediction of Skin Penetration Using Artificial Neural Network (ANN) Modeling	8
Lee et al. [27], Skin Permeability of Various Drugs with Different Lipophilicity	8
Hatanaka et al. [28], Prediction of Skin Permeability of Drugs. I. Comparison with Artificial Membrane	8
Ghosh et al.[25], Comparison of skin Permeability of drugs in mice and human cadaver skin	4
Wiechers [29], The barrier function of the skin in relation to percutaneous absorption of drugs	4
Martinez-Pla et al. [30], Evaluation of the pH effect of formulations on the skin permeability of drugs by biopartitioning micellar chromatography	4
Lim et al. [23], Prediction of Human Skin Permeability Using a Combination of Molecular Orbital Calculations and Artificial Neural Network	11
Lipinski et al. [24], Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings	37
O'Neil, M.J. (ed.).[31] The Merck Index - An Encyclopedia of Chemicals, Drugs, and Biologicals	6
Wilschut et al. [26], Estimating skin permeation. The validation of five mathematical skin permeation models	99
Online databases DrugBank [32], LookChem [33] and ChemichalBook [34]	211

Development of an accurate skin permeability model is becoming increasingly important as skin has been more utilized in recent development of drug delivery methods [14].

2 Methods

2.1 Identification of parameters for prediction of drug permeability

The database used in this study contains information about 3 parameters for 400 different drugs. Those parameters are molecular weight, partition coefficient (logP) and melting temperature. This database is separated into groups based on their activity.

Multiple studies were conducted, aiming to determine skin permeability basing on different parameters. Study by Lim et al. [23] was developed to predict skin permeability coefficient of compounds from their three-dimensional molecular structure using a combination of molecular orbital calculation and artificial neural network [24]. Based on Moriguchi log(P) calculation Lipinski et al. [24] showed that poor absorption or permeation is more likely when there are more than 5 H-bond donors, 10 H-bond acceptors, and that the molecular weight s greater than 500 and the calculated log(P) (Clog(P)) is greater than 5 (or Mlog(P).4.15). hosh et al. [25] compared skin permeability of drugs in mice and human cadaver skin and found difference and similarity in human and mice permeability. Number of 99 different chemicals and their physiochemical properties were collected from Wilschut et al. [26] study of validation of five mathematical skin permeation model.

The data form such studies were collected for establishing the database for development of ANN in this study. Table 1 presents studies, drug samples and number of drug samples taken from each study for developed database.

Not all parameters available in each study were taken when developing database for development of ANN for skin permeability prediction. Only information about three parameters were used, and those are molecular weight, partition coefficient (logP) and melting temperature. Also, online available databases were used for acquiring needed information such as DrugBank [32], LookChem [33] and ChemichalBook [34].

2.2 Artificial Neural Network for prediction of drug permeability

In this paper, system for prediction of permeability of drug based on pharmacokinetic parameters was developed. The neural network is developed based on three input parameters as follows: molecular weight, partition coefficient log(P), and melting temperature.

Network type used in this paper for solving the problem of prediction of skin permeability, was feedforward, back-propagation architecture, commonly used for solving this type of problems [35-40]. The typical back-propagation network has an input layer- which is consist only of network inputs. It is then followed by a hidden layer which consists of number of neurons, or hidden units which are placed by parallel [41,42]. The network output is also formed by weighted summation that consists of outputs of the neurons in the hidden layer. As it can be seen from Fig. 1, network inputs are not affected with network output in any way. The output is result of modifiable synapses, represented as summations of signals from hidden and input layer.

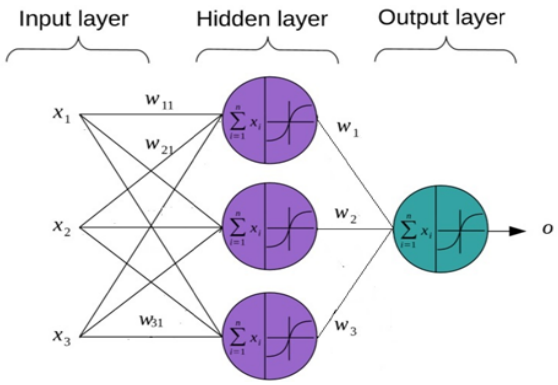

Fig. 1. Architecture of a single layer feedforward neural network

Computationally, learning in this type of ANN architecture is regression process of adapting weights and biases until minimum error value is achieved [41]. Simple iterative algorithm that can be used in training of this type of neural network is back-propagation algorithm. The output values are compared with the correct answer to compute the value of some predefined error-function. The weights and biases of neural network are corrected in each iteration. Network performance function that is used is Mean Squared Error.

For training Levenberg – Marquardt algorithm (LMA) is used, which is common training algorithm in data prediction [36]. Training is performed using training and estimation dataset of 300 (75% of total dataset) samples. The reason

for choosing this number of samples is to achieve appropriately balanced training dataset. This dataset ratio was determined by "trial and miss" method while assessing the ANN performance. The starting network weights were initialized according to learning algorithm used.

2.3 Ensemble techniques based on Decision Trees for prediction of skin permeability

Regression techniques predict continuous responses—for example, changes in temperature or fluctuations in power demand. It builds models to predict continuous data. With this information, you can make predictions about future data points. Since our data has continuous response as the output parameter, regression techniques were used.

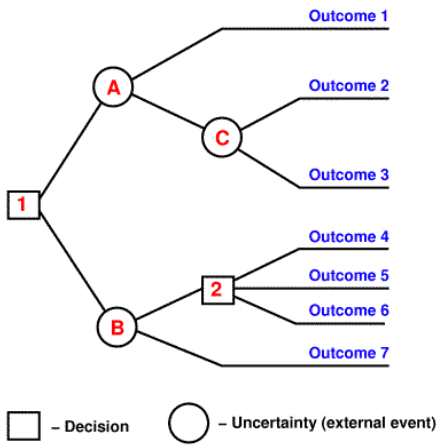

Fig. 2. Decision Tree Outline

Among many regression methods REPTree, Bagging and Random SubSpace algorithms were used in the analysis using Weka software. Weka is an open source software issued under the GNU General Public License and it contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization. It is also well-suited for developing new machine learning schemes [43].

Reduced Error Pruning Tree (REPTree). REP-Tree is a method that produces a regression tree and was designed particularly to be fast, which may however lead to some reduction in its predictive accuracy, by comparison with other decision tree methods. RepTree uses the regression tree logic and creates multiple trees in different iterations. After that it selects best one from all generated trees. That will be considered as the representative M5-Rules is an M5P version that produces a model consisting of a list of *If-Then* regression rules, rather than regression trees [44].

Bagging. The basic idea underlying Bootstrap Aggregating (Bagging) which was first introduced by Breiman [45] is the recognition that part of the output error in a single regression tree is due to the specific choice of the training data set. Therefore, if several similar data sets are created by resampling with replacement (that is, bootstrapping) and regression trees are grown without pruning and averaged, the variance component of the output error is reduced votes classifiers generated by different bootstrap samples (replicates). REPTree was used as a regression algorithm.

Random SubSpace. Random subspace method [46] also called attribute bagging or feature bagging, is an ensemble learning method that attempts to reduce the correlation between estimators in an ensemble by training them on random samples of features instead of the entire feature set. It is an ensemble learning method that attempts to reduce the correlation between estimators in an ensemble by training them on random samples of features instead of the entire feature set. REPTree was used for regression.

Table 2. Distribution of dataset in Artificial Neural Network

Total number of samples included in dataset based on research **445**	**Number of samples**	**Percentage of samples**
No. Of sample outliers excluded from the dataset	45	10.11%
Total number of samples used for ANN development **400**	**Number of samples**	**Percentage of samples**
No. Of samples for training	240	60%
No. Of samples for validation	60	15%
No. Of samples for testing	100	25%

2.4 Dataset Distribution

The squaring of the errors tends to heavily weight statistical outliers, affecting the accuracy of the results. Outlier omission was performed by applying regression analysis of ANN and Bagging algorithms on full database with the expected results. Using this principle, outliers with 0.2 error rate or more were later excluded from the dataset which were in total 45 as it can be seen from Table 2.

3 Results

In this study, impact of number of neurons in hidden layer, training functions as well as dataset distribution during training on ANN output coefficient correlation was examined. Based on the ANN performance, the optimal architecture was chosen for solving problem of prediction of skin permeability. Performance of chosen ANN architecture is compared to three different regression algorithms: REPTree, Bagging, Random SubSpaceDeveloped

3.1 Predictive ability of neural network architecture with different number of neurons in hidden layer

The ANN was trained with 10, 20, 50 and 100 neurons in the hidden layer with *trainlm* (Levenberg-Marquardt backpropagation) training function with mean squared error as a performance function. This function that weight and bias values according to Levenberg-Marquardt optimization. It is often the fastest backpropagation algorithm, and is highly recommended as a first-choice supervised algorithm, although it does require more memory than other algorithms [36].

As it can be seen from Fig. 3, ANN architecture with 20 neurons gave the best results regarding ANN performance so this number of neurons was chosen for system development.

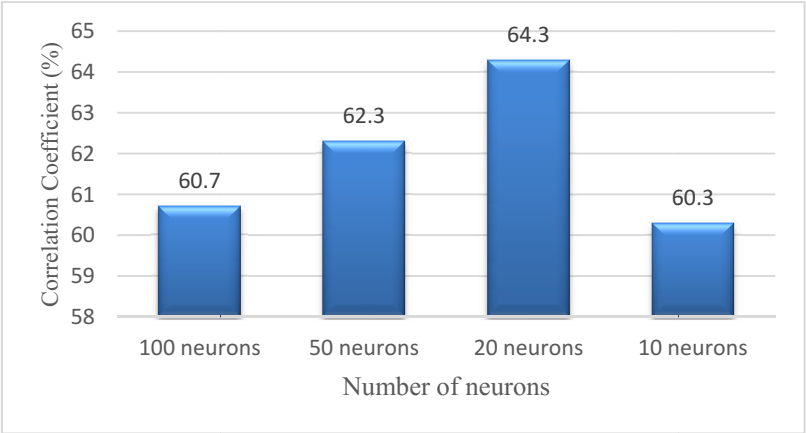

Fig. 3. Representation of different number of neurons in hidden layer

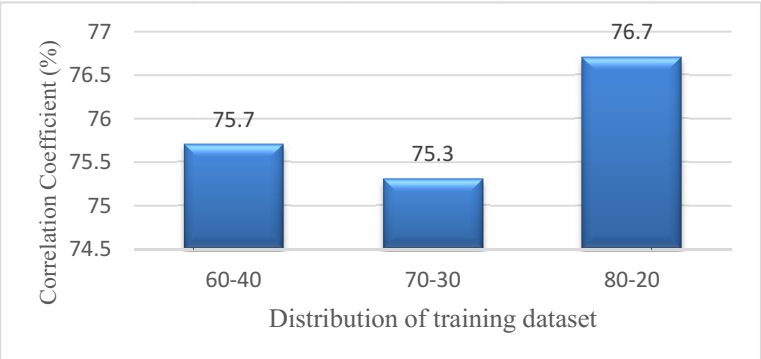

Fig. 4. Representation of different distribution of training dataset

Table 3. Training and testing of ANN for prediction of drug permeability

Dataset used for development of ANN		
No. of samples	Accurate prediction	False prediction
300		
Training		
240	92.36	7.64
Validation		
60	88.79	11.21
Testing of ANN		
No. of samples	Accurate prediction	False prediction
100	76.7%	23.3%

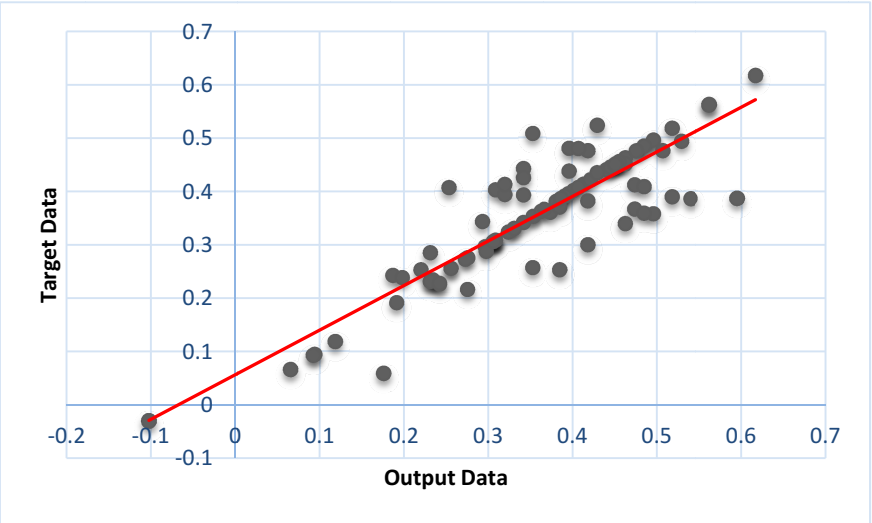

Fig. 5. Scatter plot of ANN performance

3.2 Predictive ability of neural network architecture with different distribution of dataset

The network with 20 neurons, was trained with 300 samples assessing the impact of different ratio of training dataset. Since, experts [38, 39] recommend using 70%-30% training dataset division for smaller datasets, different data ratio were inspected to see if this dataset division fits this case. Recommended dataset division was compared to 60%-40%, and 80%-20%. The results of training are presented in Fig.4. As it can be seen from the figure, ANN with 80%-20% training dataset division provided the best results and was chosen for solving problem of skin permeability prediction.

3.3 Predictive ability of neural network for prediction of drug permeability

Feedforward two-layer neural network architecture with 20 neurons in hidden layer was validated with 100 samples, or 15% of samples from the whole dataset after excluding outliers. Percentages of hits and misses of the artificial neural network in prediction of drug permeability in this study are presented in Table 3.

As it can be seen from Table 3, from the total sum of 100 samples, the ANN correctly predicted 77 samples or 76.7%, while incorrectly predicted 23 samples or 23.3% of testing dataset. The ANN performance during training is shown in Fig.5 where output data presents the value of skin permea-

bility calculated by ANN and target data represents the actual value of skin permeability.

In respect to studies by Agatonovic-Kustrin et al.[16] and Cheng et al. [17] who have used ANNs in determining the quantitative structure-permeability relationship of penetration across skin or polydimethylsiloxane membranes as the model of skin permeation, skin permeability in this paper is determined based on three parameters which increases accuracy of prediction. Degim et al. [18] analyzed skin permeability of 40 compounds by ANN, while in this paper dataset of 400 samples was used for ANN development.

3.4 Predictive ability of REPTree, Bagging, and Random SubSpace

The data distribution is the same as shown in Table 2. The reason behind it is to see how these algorithms compare against the ANN in prediction in same conditions.

Table 4. Ensemble Techniques performance

Algorithm	Full Dataset	Without Outliers
REPTree	65.9%	79.3%
Bagging	69.9%	81%
Random SubSpace	70.1%	73.3%

3.5 Comparison of ANN and Ensemble Techniques

As it can be seen in Table 5, the results of ensemble techniques outperform ANN's results. Random SubSpace gives the best result for the full dataset (70.1%), while bagging outperforms the rest (81%). The reason behind this can be the insufficient number of samples or input parameters for ANN. Compared to previous research we can see that number of samples used here is smaller. Whereas number of samples in Badran et al.[47] and Jerez-Aragonés et al. [48] are 18 years of forecasting data and 1035 patients, respectively, the size of our data was significantly lower. Furthermore, input parameters in Jerez-Aragonés et al. [48] was 85, while we had 3.

Table 5. ANN and Ensemble Techniques performance

Algorithm	Full Dataset	Without Outliers
ANN	64.3%	76.7%
Bagging	69.9%	**81%**
Random SubSpace	**70.1%**	73.3%

Interesting area of research in machine learning and neural networks are combining a set of learned models to improve classification and regression estimates. The goal of combining learned models is to obtain a more accurate prediction than can be obtained from any single source alone. Badran et al. [47] published a paper that combined artificial neural network and regression modeling methods to predict electrical load. Similar combination of neural network and decision trees model were done by Aragonés et al. [48] in their prognosis of breast cancer.

4 Conclusion

This paper presents development of Artificial Neural Network (ANN) for prediction of skin permeability based on following parameters of drugs: molecular weight, partition coefficient (logP) and melting temperature. The ANN performance was compared to performance of three different Ensemble Techniques based on three models.

Feedforward neural network architecture with 2 layers and 20 neurons in hidden layer was chosen for solving this problem. Neural network was trained with 300 samples and developed system was tested with 100 samples. The dataset was made based on previous research in this area, and before the dataset was used for ANN development, outliners were left out of the dataset based on mean square error performance.

Developed ANN gave regression result of 0.76, which compared to other researches in this area can be described as good result, taken into consideration the small amount of input attributes (3).

For Ensemble Techniques REPTree, Bagging and Random SubSpace were implemented. Implementation of Ensemble Techniques as a regression tool tend to yield better results when there is a significant diversity of the model in the output layer. For that reason, these techniques were significant in this research.

Using ANN for solving this problem though has more advantages over ensemble techniques, given that neural networks are already covered by mathematical models and usage of more complex computer technologies are not needed. Even though ANN is an extremely useful tool, it underperformed to Bagging and Random SubSpace by 4.3%, due to the low number of input parameters. Interconnection between tested algorithms and pharmacy can improve this field of science and even try to become one for itself.

Approach by means of correlation between skin permeability and parameters used in this research can be useful in

detecting contribution of each parameter in order to determine permeability. One of the precedence of this research is that unlike majority of former ANN developed models, it didn't require experimental analysis which gave us more time for construction of drug base and development of better ANN model. This can lead to less complex calculations and faster predictions. Clear benefit is that this ANN model can provide precise prediction of skin permeability for drugs that are still in stage of development. Next to this, it is safe to say that it can also serve as reducing agent of toxicity and help in upcoming therapies for different diseases.

Simulation, modeling, as well as analysis of different databases are these algorithms' characteristics allowing it to be highly used in today's research in every scientific field, including pharmacy. The expectations are immensely huge and they all rely on fact that we can gain more accurate results using these techniques rather than conventional techniques. This tells a lot about effort given for development and improvement of ANN and other techniques together with all fields of science, especially pharmaceutical concepts.

For further research, authors plan on using this system in laboratory conditions and investigate possibility of expanding the number of input parameters to increase system performance.

References

1. Kolarsick, P. A., Kolarsick, M. A., & Goodwin, C. (2011). Anatomy and physiology of the skin. *Journal of the Dermatology Nurses' Association*, *3*(4), 203-213.
2. Saini, S., Singh, S. K., Garg, A., Khanna, K., Shandil, A., & Mishra, D. N. (2010). Prediction of skin penetration using artificial neural network. *Int. J. Eng. Sci. Technol*, *2*, 1526-1531.
3. Bolzinger, M. A., Briançon, S., Pelletier, J., & Chevalier, Y. (2012). Penetration of drugs through skin, a complex rate-controlling membrane. *Current Opinion in Colloid & Interface Science*, *17*(3), 156-165.
4. Kielhorn, J., Melching-Kollmuss, S., & Mangelsdorf, I. (2006). Environmental Health Criteria 235: Dermal Absorption. *World Health Organization, Geneva, Switzerland*.
5. Schaefer, H., Schalla, W., Zesch, A., & Stüttgen, G. (2013). *Skin permeability*. Springer Science & Business Media.
6. Scheuplein, R. J. (1965). Mechanism of Percutaneous Adsorption: I. Routes of Penetration and the Influence of Solubility. *Journal of investigative Dermatology*, *45*(5), 334-346.
7. Scheuplein, R. J. (1967). Mechanism of Percutaneous Absorption: II. Transient Diffusion and the Relative Importance of Various Routes of Skin Penetration** From the Research Laboratories of the Department of Dermatology of the Harvard Medical School at the Massachusetts General Hospital, Boston, Massachusetts 02114. *Journal of Investigative Dermatology*, *48*(1), 79-88.
8. Scheuplein, R. J., & Blank, I. H. (1971). Permeability of the skin. *Physiol Rev*, *51*(4), 702-747.
9. Scheuplein, R. J., & Blank, I. H. (1973). Mechanism of percutaneous absorption. IV. Penetration of nonelectrolytes (alcohols) from aqueous solutions and from pure liquids. *Journal of Investigative Dermatology*, *60*(5), 286-296.
10. Flynn, G. L. (1990). Physicochemical determinants of skin absorption. *Principles of route-to-route extrapolation for risk assessment*, 93-127.
11. Swarbrick, J., & Boylan, J. C. (Eds.). (2000). *Encyclopedia of Pharmaceutical Technology: Volume 20-Supplement 3*. CRC Press.
12. Flynn, G. L., & Stewart, B. (1988). Percutaneous drug penetration: choosing candidates for transdermal development. *Drug Development Research*, *13*(2-3), 169-185.
13. Assessment, D. E. (1992). Principles and Applications. *Office of Research and Development, Washington, DC*, 48-49.
14. Sutariya, V., Groshev, A., Sadana, P., Bhatia, D., & Pathak, Y. (2013). Artificial neural network in drug delivery and pharmaceutical research. *The Open Bioinformatics Journal*, *7*, 49-62.
15. Agatonovic-Kustrin, S., & Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *Journal of pharmaceutical and biomedical analysis*, *22*(5), 717-727.

16. Agatonovic-Kustrin, S., Beresford, R., & Yusof, A. P. M. (2001). ANN modeling of the penetration across a polydimethylsiloxane membrane from theoretically derived molecular descriptors. *Journal of pharmaceutical and biomedical analysis*, *26*(2), 241-254.
17. Chen, L. J., Lian, G. P., & Han, L. J. (2007). Prediction of human skin permeability using artificial neural network (ANN) modeling. *Acta Pharmacologica Sinica*, *28*(4), 591.
18. Değim, T., Hadgraft, J., İlbasmiş, S., & Özkan, Y. (2003). Prediction of skin penetration using artificial neural network (ANN) modeling. *Journal of pharmaceutical sciences*, *92*(3), 656-664.
19. Pugh, W. J., Degim, I. T., & Hadgraft, J. (2000). Epidermal permeability–penetrant structure relationships: 4, QSAR of permeant diffusion across human stratum corneum in terms of molecular weight, H-bonding and electronic charge. *International journal of pharmaceutics*, *197*(1), 203-211.
20. Geurts, P., Irrthum, A., & Wehenkel, L. (2009). Supervised learning with decision tree-based methods in computational and systems biology. *Molecular Biosystems*, *5*(12), 1593-1605.
21. Lemon, S. C., Roy, J., Clark, M. A., Friedmann, P. D., & Rakowski, W. (2003). Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. *Annals of behavioral medicine*, *26*(3), 172-181.
22. Hu, Y. J., Ku, T. H., Jan, R. H., Wang, K., Tseng, Y. C., & Yang, S. F. (2012). Decision tree-based learning to predict patient controlled analgesia consumption and readjustment. *BMC medical informatics and decision making*, *12*(1), 1.
23. Lim, C. W., Fujiwara, S. I., Yamashita, F., & Hashida, M. (2002). Prediction of human skin permeability using a combination of molecular orbital calculations and artificial neural network. *Biological and Pharmaceutical Bulletin*, *25*(3), 361-366.
24. Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2012). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced drug delivery reviews*, *64*, 4-17.
25. Ghosh, B., Reddy, L. H., Kulkarni, R. V., & Khanam, J. (2000). Comparison of skin permeability of drugs in mice and human cadaver skin. *Indian journal of experimental biology*, *38*(1), 42-45.
26. Wilschut, A., Wil, F., Robinson, P. J., & McKone, T. E. (1995). Estimating skin permeation. The validation of five mathematical skin permeation models. *Chemosphere*, *30*(7), 1275-1296.
27. Lee, C. K., Uchida, T., Kitagawa, K., Yagi, A., Kim, N. S., & Goto, S. (1994). Skin permeability of various drugs with different lipophilicity. *Journal of pharmaceutical sciences*, *83*(4), 562-565.
28. Tomomi, H., Masami, I., Kenji, S., & Yasunori, M. (1992). Prediction of skin permeability of drugs. II. Development of composite membrane as a skin alternative. *International journal of pharmaceutics*, *79*(1-3), 21-28.
29. Wiechers, J. W. (1989). The barrier function of the skin in relation to percutaneous absorption of drugs. *Pharmaceutisch Weekblad*, *11*(6), 185-198.
30. Martínez-Pla, J. J., Martín-Biosca, Y., Sagrado, S., Villanueva-Camanas, R. M., & Medina-Hernandez, M. J. (2004). Evaluation of the pH effect of formulations on the skin permeability of drugs by biopartitioning micellar chromatography. *Journal of Chromatography A*, *1047*(2), 255-262.
31. Neil, M. J. (2006). The Merck Index: An Encyclopedia of chemicals, drugs and biologicals. Whitehouse station, New Jersey: Published by Merck Research Laboratories, Division of Merck and Co.
32. DrugBank, A. (2016). Open data drug & drug target database. *Accessed online July*, *12*.
33. LookChem, (2016). *Accessed online, July, 12.*
34. Chemicalbook (2016). Chemical Search Engine. *Accessed online, July, 12.*
35. Mustafidah, H., Hartati, S., Wardoyo, R., & Harjoko, A. (2014). Selection of Most Appropriate Backpropagation Training Algorithm in Data Pattern Recognition. *arXiv preprint arXiv:1409.4727*.

36. Fojnica, A., Osmanović, A., & Badnjević, A. (2016). Dynamical Model of Tuberculosis-Multiple Strain Prediction based on Artificial Neural Network. . In *Embedded Computing (MECO), 2016 5th Mediterranean Conference on* (pp. 290-294). IEEE.
37. Alić, B., Sejdinović, D., Gurbeta, L., & Badnjevic, A. (2016, June). Classification of stress recognition using Artificial Neural Network. In *Embedded Computing (MECO), 2016 5th Mediterranean Conference on* (pp. 297-300). IEEE.
38. Aljović, A., Badnjević, A., & Gurbeta, L. (2016, June). Artificial neural networks in the discrimination of Alzheimer's disease using biomarkers data. In *Embedded Computing (MECO), 2016 5th Mediterranean Conference on* (pp. 286-289). IEEE.
39. Badnjević, A., Gurbeta, L., Cifrek, M., & Marjanovic, D. (2016). Classification of Asthma Using Artificial Neural Network. In *39th International convention on information and communication technology, electronics and microelectronics (MIPRO)*. IEEE.
40. Badnjević, A., & Gurbeta, L. (2016). Development and Perspectives of Biomedical Engineering in South East European Countries. In *39th International convention on information and communication technology, electronics and microelectronics (MIPRO)*. IEEE.
41. Halilović, S., Avdihodžić, H., & Gurbeta, L. (2016, June). Micro cell culture analog apparatus (μCCA) output prediction using Artificial Neural Network. In *Embedded Computing (MECO), 2016 5th Mediterranean Conference on* (pp. 294-296). IEEE.
42. Chen, H., Xiong, S., & Ren, X. (2014, May). Evaluating the Risk of Metabolic Syndrome Based on an Artificial Intelligence Model. In *Abstract and Applied Analysis* (Vol. 2014). Hindawi Publishing Corporation.
43. Witten, I. H., & Frank, E. (2005). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.
44. Holmes, G., Hall, M., & Prank, E. (1999, December). Generating rule sets from model trees. In *Australasian Joint Conference on Artificial Intelligence* (pp. 1-12). Springer Berlin Heidelberg.
45. Breiman, L. (1996). Bagging predictors. *Machine learning*, *24*(2), 123-140.
46. Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE transactions on pattern analysis and machine intelligence*, *20*(8), 832-844.
47. Badran, S. M., & Abouelatta, O. B. (2011). Forecasting Electrical Load using ANN Combined with Multiple Regression Method. *The Research Bulletin of Jordan ACM*, *2*(2), 152-158.
48. Jerez-Aragonés, J. M., Gómez-Ruiz, J. A., Ramos-Jiménez, G., Muñoz-Pérez, J., & Alba-Conejo, E. (2003). A combined neural network and decision trees model for prognosis of breast cancer relapse. *Artificial intelligence in medicine*, *27*(1), 45-63.

Corresponding author:

Author: Ahmed Osmanović
Institute: International Burch University
Street: Nerkeza Smajlagića 1
City: Sarajevo
Country: Bosna and Herzegovina
Email: ahmed_osmanovic@hotmail.com