**PROCEEDINGS OF IKSTC 2024 - 3rd INTERNATIONAL KARATEKIN SCIENCE AND TECHNOLOGY CONFERENCE **

**21-22 NOVEMBER 2024, ÇANKIRI, TURKEY**

STATISTICAL STUDIES IN THE PHARMACEUTICAL INDUSTRY

***Sebnem SENOL 1,*** ***, Aynur DEMİR1*** ***, Emel AKYOL 1***

*1 Institute of Natural Sciences, Faculty of Chemistry-Metallurgy, Chemical Engineering, Yildiz Technical University, Istanbul, Turkey*

|  |
| --- |
| **Abstract**  In this study, the data from the article "Synthesis and characterization of hydrogels based on poly(2-hydroxyethyl methacrylate) for drug delivery under UV irradiation" were used within the scope of statistical study in the pharmaceutical industry. Two statistical methods called "Artificial Neural Networks (ANN)" in the Matlab programme and "Response Surface Method (RSM)" in the Minitab package programme were used. It was determined how close the results obtained in the programme are to the results found in the laboratory experiment. Today, with the developing technology, both time and cost are becoming increasingly valuable. For this reason, it is aimed to find results close to reality by saving time and cost with statistical methods. |
| ***Keywords:*** *Artificial Neural Networks, Response Surface Method, Matlab, Minitab, Drug release* |

# Introduction

With the development of technology, prediction modeling is performed to solve complex structures in some areas for the future. "Artificial Neural Networks (ANN)" is one of the methods that make predictions by using artificial intelligence [1]. "Artificial Neural Networks (ANN)", which was discovered by being inspired by the human brain, is a computer system and prediction method that serves this purpose in addition to the fact that it performs many abilities such as producing, designing and predicting information just like our brain without any help [2]. Artificial Neural Networks perform non-linear modelling between input and output parameters without the need for prior knowledge and without making a prediction. Inputs are introduced to the network and then output variables are given in response to the inputs. The network is trained by learning the relationship between input and output. The reason why this method is preferred is learning by teaching [3].

The artificial nerve cell consists of five parts;

* + Input variables
  + Weights
  + Collection function
  + Activation function
  + Output variables [4]

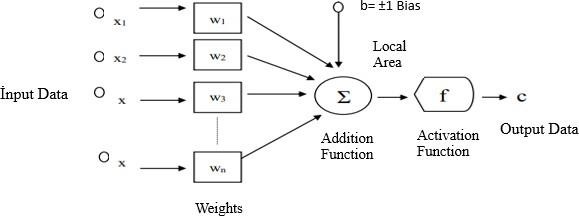


Figure 1. Working Principle of Artificial Neural Network [4]

\* Corresponding author. *e-mail address:* [*sebnemsenol\_@hotmail.com*](mailto:sebnemsenol_@hotmail.com)

The sum function is a type of function that calculates the effect of all inputs and weights on this processing element. The sum function calculates the net input to the cell. If all of the net input collected in the cell is,

𝒏𝒆𝒕 = ∑𝒏 𝐰**ij xi *+ b***

𝒊=𝟏

is obtained as. In this equation xi : input value of neuron i, wij : weight coefficients,

n : the total number of inputs to a cell,

b is the threshold value and Σ is the sum function [5].



Network architecture is selected and structure features are determined

Characteristic features of neuron functions are determined

Learning algorithm is selected and parameters are determined

Training and testing data are created

The network is trained and tested

Figure 2: Artificial Neural Networks Workflow

The Response Surface Method is a technique that combines mathematics and statistics, which is used in analysis and modeling in some applications where the answers sought are affected by different variables and these answers are aimed to reach optimized conditions. The response surface method, which is used in industry and many other fields, is frequently used in scientific research as it enables us to reach the most appropriate answer with fewer experiments [6]. The response surface method consists of 3 stages: model validation, mathematical modeling and design of experiments. With the design of the experiment, it is possible to establish a relationship between the dependent and independent parameters and to determine which parameter is the most important. Polynomial equations created in the response surface method are widely used because they provide ease in estimating the response surface function. To see to what extent the estimated model represents the real values and how accurate it is, the model should be validated [7].

# Materials and Methods

In the present study, the article "Synthesis of HEMA (Poly 2-hydroxyethylmethacrylate) based hydrogels for controlled drug release under the influence of photoinitiators"[8] was used. In this article, HEMA-based hydrogels were prepared with photoinitiators (Irg 184, Irg 651, Irg 2959) and crosslinker (ethylene glycol dimethacrylate). Donepezil HCL, HAp or TiO2 , the active ingredient of the drug used in the treatment of Alzheimer's disease, was added to HEMA-based hydrogels by photopolymerization method and the swelling and release effects of hydrogels were investigated. Release analyses were performed in different pH environments and the most suitable release environment was determined. It was determined that the synthesized hydrogels can be used in the release of Donepezil HCl drug active substance. The results [8] were predicted by Artificial Neural Network (ANN) and Response Surface Method (RSM). For Artificial Neural Networks (ANN), Matlab programme R2018b version Neural Network Toolbox was used. For the Response Surface Method (RSM), the Minitab package programme Design Of Experiment (DOE) was used. The prediction results were statistically compared with the experimental results.

# Result and Discussion

Firstly, predictions were made with Artificial Neural Network and compared with the results in the paper.

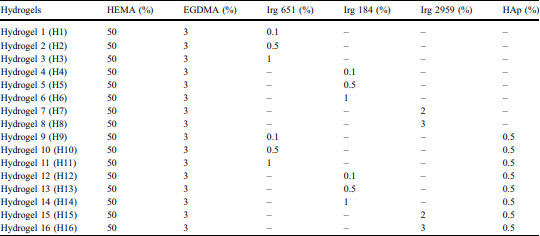


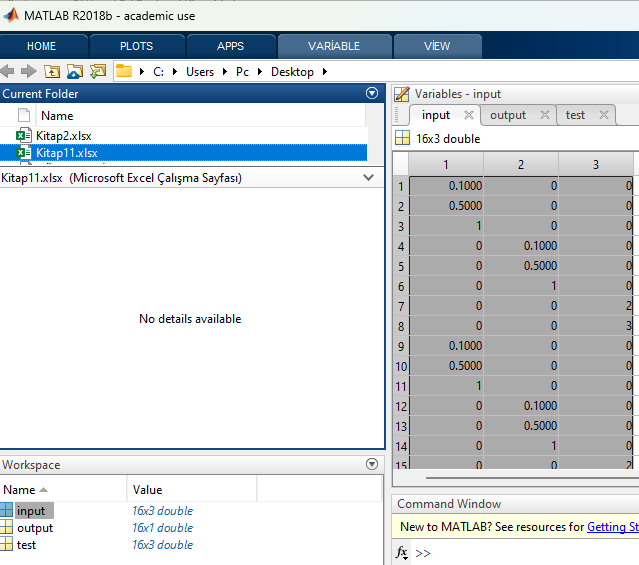
Table 1. Formulation of HEMA-based hydrogels

HEMA-based hydrogels were synthesized using ethylene glycol dimethacrylate crosslinker and Irgacure 184, Irgacure 651, Irgacure 2959 photoinitiators and HAp as additive. The following table shows the hydrogels synthesized at different times according to pH 1.2 taken from the article "Synthesis of HEMA (Poly 2-hydroxyethyl methacrylate) based hydrogels for controlled drug release under the influence of photoinitiators"[8]. The detected results were predicted by Artificial Neural Network and the results were statistically compared.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **PH 1.2** | | | | | | | | | | | | | | | | |
| Time(min) | H1 | H2 | H3 | H4 | H5 | H6 | H7 | H8 | H9 | H10 | H11 | H12 | H13 | H14 | H15 | H16 |
| 10 | 14 | 18 | 30 | 12 | 13 | 14 | 27 | 53 | 6 | 19 | 34 | 3 | 12 | 14 | 26 | 36 |
| 20 | 15 | 27 | 41 | 14 | 16 | 20 | 32 | 69 | 21 | 34 | 49 | 11 | 14 | 15 | 30 | 45 |
| 30 | 22 | 28 | 44 | 15 | 22 | 25 | 36 | 73 | 32 | 37 | 61 | 23 | 30 | 31 | 35 | 50 |
| 60 | 39 | 44 | 70 | 20 | 24 | 26 | 46 | 83 | 33 | 54 | 87 | 29 | 39 | 40 | 40 | 56 |
| 120 | 41 | 65 | 80 | 32 | 34 | 37 | 58 | 94 | 41 | 80 | 87 | 42 | 44 | 45 | 55 | 75 |
| 180 | 70 | 84 | 93 | 43 | 43 | 52 | 67 | 94 | 49 | 80 | 87 | 45 | 46 | 47 | 71 | 89 |
| 240 | 74 | 91 | 93 | 48 | 49 | 53 | 68 | 94 | 49 | 80 | 87 | 46 | 48 | 50 | 71 | 89 |
| 300 | 74 | 91 | 93 | 50 | 51 | 53 | 72 | 94 | 52 | 80 | 87 | 47 | 50 | 51 | 71 | 89 |
| **360** | **74** | **91** | **93** | **50** | **51** | **53** | **72** | **94** | **70** | **80** | **87** | **49** | **50** | **51** | **71** | **89** |

Table 2: Drug Release Value at pH 1.2

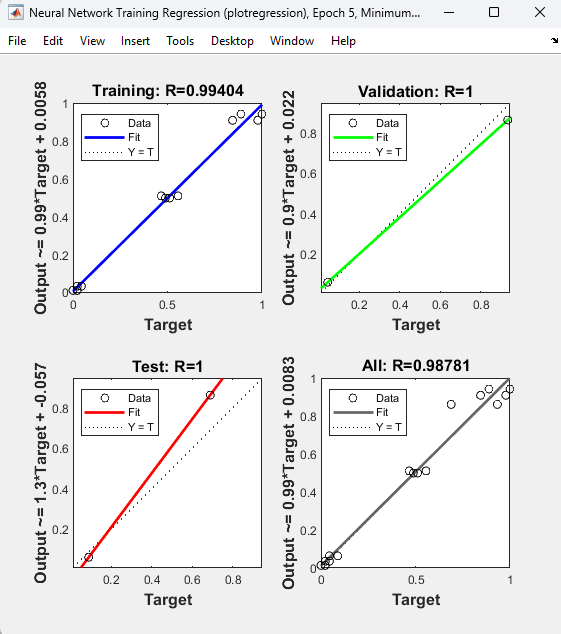
MATLAB programme was used in the development of the ANN model. With the help of the codes written in the MATLAB programme, the data set was introduced to the network and training was performed. The most appropriate network architecture was determined and the model was created. In the study, a feed-forward network structure was used to predict the drug release rate of the hydrogel. The first step of the prediction process in artificial neural networks starts with the selection of input, output and test data. Input data consists of 3 columns as Irg 651, Irg 184, Irg 2959. The output data are synthesized hydrogels. The number of neurons in the input layer of the network is 3, the number of neurons in the intermediate layer is 10 and the number of neurons in the output layer is 1. The Sigmoid transfer function was used in the hidden (intermediate) and output layers. The number of hidden layers and the number of neurons in the layers were determined by trial and error. Levenberg-Marquardt optimization method was chosen for training the network and adjusting the weights. A total of 16 data were obtained as a result of the experimental studies. In the table given above, separate predictions were made in Artificial Neural Networks for pH 1.2 in the time period from 10 minutes to 360 minutes, and the most efficient result was obtained for 360 minutes. The prediction study at pH 1.2 at 360 minutes is shown in the image below. The data were normalized before being introduced to the network.The normalization process is applied to increase model success. Firstly, the data were transferred to the Matlab programme. Input variables are assigned to the input section, output variables are assigned to the output section and variables to be tested are assigned to the test section. The 16 variables from H1 to H16 in Table 2 were transferred as both input and test data.

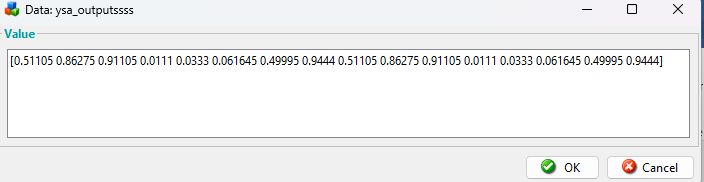
Figure 3: Data transfer screen in ANN

The resulting artificial neural network model accordingto the selections on the screen above is shown in Figure4.

Figure 4: Structure of ANN



The training for the test data in ANN was completed and the prediction processes were performed and the results are as follows.



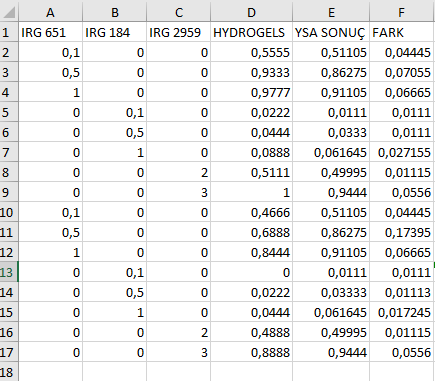
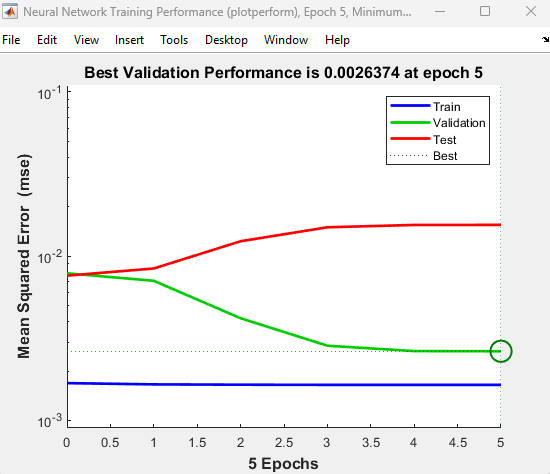
Figure 7: ANN result screen

Figure 5. Regression graph in ANN

The training performance of the selected model is given in the graph in Figure 6.

Table 3: ANN training result screen

The regression equation resulting from the experimental result values and the results generated by the Artificial Neural Network is as follows



1,2

1

0,8

0,6

0,4

0,2

0

y = 0,997 - 0,011

R² = 0 728

0 0 , 2 0 , 4 0 , 6 0 , 8 1

ANN PREDİCTED VALUE

EXPERİMENTAL VALUE

Figure 6. ANN training performance

Figure 8. Release graph of predicted values and actual values

After the artificial neural network study, the study with the Minitab response surface method (RSM) is as follows. As a first step, the Minitab programme is opened, central composite is selected as the design type from the create response surface design option, the number of factors is selected as 3 for our 3 variables[8] and a design is created.

After selecting the central composite design type, input and output variables are assigned.

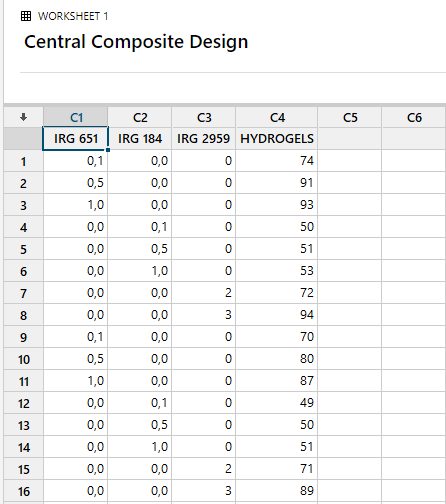
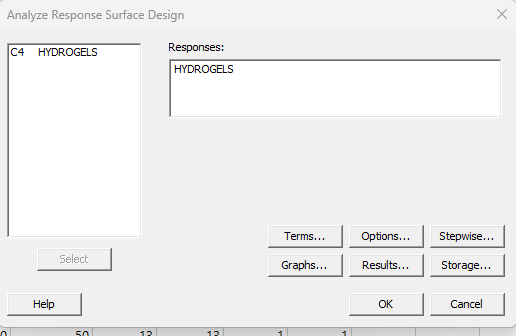
 

Figure 9. Data transfer screen in Minitab

Input and output variables are entered separately.

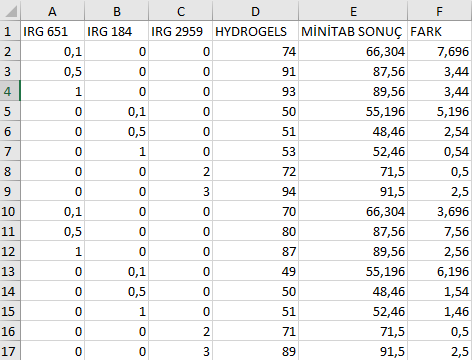
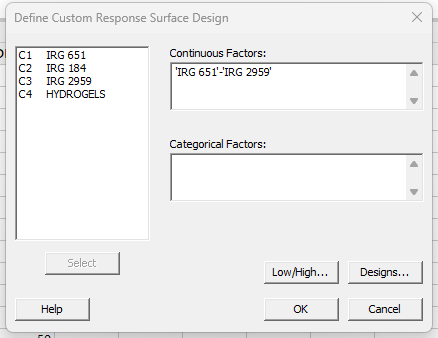


Table 4. Analysis results

Figure 10. Entering input variables Figure 11. Entering output variable

The results of the analysis are as follows.



Figure 12. R-value results



Figure 13. Equation found as a result of the analysis

100

80

y = 0,9999x + 1E-04

R² = 0,9453

60

40

20

0

0

20

40

60

80 100

MİNİTAB PREDİCTİON VALUES

Figure 14. Release graph of predicted values and actualvalues

The regression equation resulting from the experimental result values and the results generated bythe Response Surface Method (RSM) is as follows

# Conclusion

When the data obtained by using Artificial Neural Networks and Response Surface Method in this studyare analyzed, the following results emerge:

In the article "Synthesis and characterization of hydrogels based on poly(2-hydroxyethyl methacrylate) for drug delivery under UV irradiation"[8], ANN gave the closest result to the data obtained as a result of the experimental study. In Figure 5, the correlation coefficient shows the relationship between outputsand targets. The values between outputs and targets are very close to 1. The correlation values of the model are 0.99404 for training, 1 for validation and 1 for testing. The overall correlation coefficient of the data set is 0.98781. The fact that the values are close to 1 for all data shows that the data obtained with ANN training and the real data are compatible with each other. Correlation coefficient values closeto 1 means that the model is successful. As the mean squared error approaches zero, it means that better results are produced, that is, it can be said that the error is lower and more successful results are obtained. The lowest MSE value for the validation set was 0.0026374. This result again shows that the model is successful. When the Minitab analysis reports are examined, it is seen that the regression equation is significant at a 5% significance level (P= 0.000<α=0.05) R2 = 94.53%. When the results of Artificial Neural Networks training and Minitab analysis are graphed separately with experimental values, R2 value is 0,9728 for ANN and 0,9453 for RSM. In this way, the results found in the programme were verified. Based on all these, the R2 value obtained in the Artificial Neural Network prediction was higher than the Response Surface Method. According to the performance values obtained in the study, it is seenthat the Artificial Neural Network model performs more effectively than the Response Surface Method.Thus, it is seen that Artificial Neural Networks have higher prediction ability and less error than the Response Surface Method. In this case, it is possible to prefer Artificial Neural Networks in cases where regression analysis assumptions are not met and the analysis cannot be performed.

# References

[1]. (Ataseven, B. (2013). Forecasting Modelling with Artificial Neural Networks. Öneri Journal, 10(39), 101-115. [2]. (Erdoğan, E., & Özyürek, H. (2012). Price Forecasting with Artificial Neural Networks. Journal of Social and Human Sciences, 4(1), 85-92.)

[3]. Yavuz, S., & Deveci, M. (2012). The effect of statistical normalization techniques on artificial neural network performance. Erciyes University Journal of Faculty of Economics and Administrative Sciences,(40), 167-187.

[4]. Damla, Y., Temiz, T., & Keskin, E. (2020). Estimation of Water Level Using Artificial Neural Network: Yalova Gökçe Dam Example. Kırklareli University Journal of Engineering and Science, 6(1),32-49

[5].Terzi, Ö. (2009). Prediction of Lake Egirdir Water Temperature with Artificial Neural Networks Method. Journal of Süleyman Demirel University Institute of Science and Technology, 10(2), 297-302.[6].Ateş, Ç., Bayraktar, B. and Bilen, M. (2018). Determination of the Optimum Conditions of a Boron Factory Wastewater Chemical Treatment Process by Response Surface Method. Gazi University Journalof Faculty of Engineering and Architecture, 33(1).)

[7]. (Albayrak, S. Optimisation of Noise Levels in Machine Tools by Response Surface Method).

[8].Senol, S., & Akyol, E. (2018). Synthesis and characterization of hydrogels based on poly (2- hydroxyethyl methacrylate) for drug delivery under UV irradiation. Journal of Materials Science, 53,14953-14963