Improved Optimization and Chemical Characterization for Urinary Metabolite Screening via pH Manipulation and Model Exploration

Sara Ortiz Ramirez[1], Dr. Grady Hanrahan Ph.D[2].
Department of Biology, California Lutheran University[1] and Department of Chemistry, California Lutheran University[2]

Introduction

Metabolic interferences and biotransformation for chemical compounds can provide insight into the toxic and carcinogenic modes of action of foreign compounds and their metabolites. [1-4]. To allow proper investigation, the development of sensitive and selective chemical instrumentation is warranted, especially as it relates to comprehensive biomarker screening. [5-7]. The importance of enhancing chemical separation efficiency and applicability is thus paramount. Here, computational modeling techniques including neural networks and swarm intelligence optimization methods have been useful [8-10]. They allow investigators to solve complicated analytical chemistry tasks allocation problems through accelerated instrument optimization and algorithmic scalability. pH manipulation and examination is crucial to the success of chemical separations and for the characterization of chemical compounds associated with deleterious human health effects.

Research Questions

How do we study chemical characterization of environmental and health-related phenolic compounds? How do we combine computational models, pH manipulation, and the study of pharmaceuticals to assess potential toxicity for human health assessment? Finally, how will theoretical models developed relate to actual experimental work?

Methods and Materials

In order to investigate and test our hypothesis we will first conduct a thorough search of the literature using scientific databases located in the university library. Next, we will focus our efforts on developing and optimizing neural network computational platforms and particle swarm optimization algorithms to aid in future instrumental studies. MATLAB software will be utilized in all computational technique development. Next, theoretical pH manipulation and examination of pharmaceuticals will be utilized to separate and characterize these toxic compounds and their metabolic pathways. In reference to reactive metabolites, Optimization studies will follow using the computational techniques described above. Concurrently, we will assess potential reactive metabolites using MATLAB software, different pH levels, and pharmaceutical studies. Once our data is collected, we will perform series of run tests using the neural net fitting application. We will run these hidden layers which examines the input and output of the algorithms. Next, relevant statistical tools will be used to create a predictive graph to assess past data with possible values of future data. Lastly, both sets of data will be utilized to estimate the R values for future outcomes. Chemical toxicity values and reactivity will be predicted.

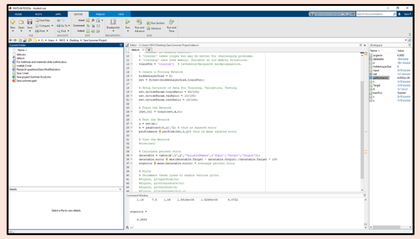

Figure 1. Shows MATLAB software used to perform a series of run tests through the neural net fitting application to collect relevant statistical data.

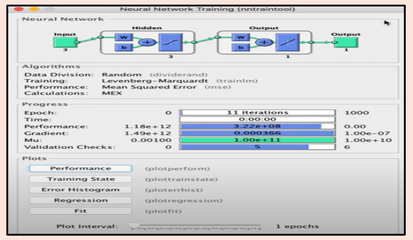

Figure 2. Shows the neural net fitting application used to collect statistical data to create a predictive graph showing the "R" value of future outcomes.

Table 1. Shows relevant statistical data for different tests performed to get a predictive graph.

60% Training/20% Validation/ 20% Testing									
Mix time (min)	Contact Volt (kV)	[Enz] (mg/mL)	Respond	Output (3 hidden layers)	%Error (3 hidden layers)	Squared Error	ouput (5 hidden layers)	% Error (5 hidden layers)	Squared Error
0.77	14.9	0.22	1246354	1176583	5.598008271	4867992441	1178341.91	5.456883821	4625644353
0.77	14.9	0.22	1163487	1176583	1.125583979	171505216	1178341.91	1.278757732	220666357.2
0.77	14.9	0.22	1176583	1176583	2.68E-11	9.59E-14	1178341.91	0.149491885	3093765.189
1.29	8.9	1.01	198151	195074	1.552856153	9467928.975	188395.2667	4.923383342	951743322.72
1.29	8.9	1.01	195074	195074	2.08E-09	1.64E-11	188395.2667	3.423692202	44605478.83
1.29	8.9	1.01	193171	195074	0.985137523	3621409.015	188395.2667	2.472282758	22807628.8
0.67	16.4	0.32	187197	187688.3333	0.262466594	241408.4444	171309.108	8.487257797	252425111.5
0.67	16.4	0.32	183459	187688.3333	2.305328893	17887260.44	171309.108	6.622674264	147619875.1
0.67	16.4	0.32	192409	187688.3333	2.453454187	22284693.78	171309.108	10.96616685	445205441.5
1.39	7.3	0.99	172714	158380.5	8.298979841	205449223.5	163127.282	5.550631657	91905163.24
1.39	7.3	0.99	157467	158380.5	0.580123522	834482.1724	163127.282	3.594583018	32038792.78
1.39	7.3	0.99	144047	158380.5	9.950571659	205449221	163127.282	13.24587256	364057162.7
0.58	15.1	0.12	205724	201934	1.842274115	14364100	201046.6487	2.273605059	21877814.91
0.58	15.1	0.12	203326	201934	0.684614855	1937664.001	201046.6487	1.121032859	5195442.217
0.58	15.1	0.12	201934	201934	9.94E-11	4.03E-14	201046.6487	0.439426382	787392.2782
1.14	7.5	1.09	155853	158073	1.424419191	4928400.187	155067.7438	0.503844161	616627.3638
1.14	7.5	1.09	156984	158073	0.693701296	1185921.092	155067.7438	1.220669776	3672037.981
1.14	7.5	1.09	159162	158073	0.684208516	1185920.908	155067.7438	2.572383007	16762934.17
Average				R= 099914	2.135651478	3.07E+08 R=0.93704		4.127613351	3.55E+08

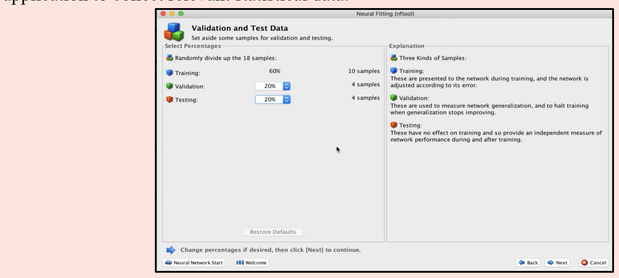

Figure 3. Shows the percentages used to measure network generalization, error during training, and independent measure of network performance during and after training of data.

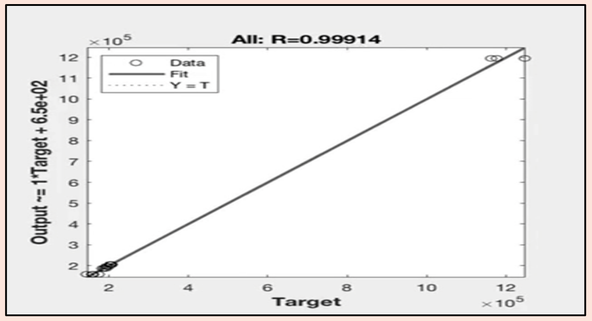

Figure 4. Shows statistical results of the "R" values and discovers patterns to draw up predictions for future outcomes.

Results

Based on the data that presented itself through the neural net fitting application, there were a series of run tests that provided statistical data to create a predictive graph. The neural net fitting application used input and output data to collect the hidden layers.

There were a series of run tests performed to collect the "R" values. This value determined the relationship between the inputs and outputs that were used to calculate the squared and percent error in the hidden layers. Based on the data tables that were created, a series of selected percentages were used to measure network generalization, error, and independent measure of network performance during and after training.

The following percentages were used:

- 60% Training/20% Validation/20% Testing
- 70%Training/15%Validation/15%Testing
- 80%Training/10%Validation/10%Testing

Each run test provided mean squared error that showed the average differences between the outputs and targets. The regression "R" values measured the correlation between the inputs and the target. An "R" value of zero was the most consistent data that showed a random relationship between the variables. The predictive graph showing the output and the target data collected provided results that analyzed predictive modeling to construct instrumental learning for future outcomes. This is expected to allow us to both separate completely and characterize complex chemical compounds associated with deleterious human health effects.

Conclusions and Discussion

During this research, there were a lot of factors that needed to be considered when attempting to create a predictive graph. One of the main factors to keep consistent is the computational error while performing the runs. Each run had to be tested multiple times to acquire efficient accuracy of the data. Any error within the tests would be due to computational errors. If there was any error in the data results a rerun was always performed. In order to get a more refined result, twelve run for each dataset were conducted to obtain a stronger "R" value. These datasets were graphed for each run performed.

The predictive graphs provided data for future outcomes using the neural net fitting application. All the runs that were performed provided an "R" value presented a strong correlation between the variables because the values were close to one. Both past and current data provided patterns that could occur in the future. The neural net fitting tool provided a measurement of network performances during and after training data to detect a linear regression. Future study of pharmaceuticals will allow for greater insight into identifying and characterizing metabolites that are crucial to environmental risk assessment.

Acknowledgments

- Dr. Allan Knox
- California Lutheran University
- Dr. Grady Hanrahan Ph.D.
- OURCS Aperture Fellowship

Contact

Sara Ortiz Ramirez
California Lutheran University.
Email: sortizramirez@callutheran.edu

References

[1]: Guo, F., Chai, L., Zhang, S., Yu, H., Liu, W., Kepp, K. P., & Ji, L. (2020). Computational Biotransformation Profile of Emerging Phenolic Pollutants by Cytochromes P450: Phenol-Coupling Mechanism. *Environmental Science & Technology, 54*(5), 2902-2912.

[2]: Borodina, Y., Sadym, A., Filimonov, D., Blinova, V., Dmitriev, A., & Poroikov, V. (2003). Predicting Biotransformation Potential from Molecular Structure. *Journal of Chemical Information and Computer Sciences, 43*(5), 1636-1646.

[3]: Guengerich, F. P. (2008). Cytochrome P450 and Chemical Toxicology. *Chemical Research in Toxicology, 21*(1), 70-83.

[4]: Baylon, J. L., Lenov, I. L., Sligar, S. G., & Tajkhorshid, E. (2013). Characterizing the Membrane-Bound State of Cytochrome P450 3A4: Structure, Depth of Insertion, and Orientation. *Journal of the American Chemical Society, 135*(23), 8542-8551.

[5]: Ji, L., & Schüürmann, G. (2015). Computational Biotransformation Profile of Paracetamol Catalyzed by Cytochrome P450. *Chemical Research in Toxicology, 28*(4), 585-596.

[6]: Koymans, L., Lenthe, J. H., Straat, R. V., Kelder, G. M., & Vermeulen, N. P. (1989). A theoretical study on the metabolic activation of paracetamol by cytochrome P-450: Indications for a uniform oxidation mechanism. *Chemical Research in Toxicology, 2*(1), 60-66.

[7]: Pavón, J. L., Sánchez, M. D., Pinto, C. G., Laespada, M. E., & Cordero, B. M. (2006). Use of Mass Spectrometry Methods as a Strategy for Detection and Determination of Residual Solvents in Pharmaceutical Products. *Analytical Chemistry, 78*(14), 4901-4908.

[8]: La, S., Yoo, H. H., & Kim, D. (2005). Liquid Chromatography−Mass Spectrometric Analysis of Urinary Metabolites and Their Pattern Recognition for the Prediction of Drug-Induced Hepatotoxicity. *Chemical Research in Toxicology, 18*(12), 1887-1896.

[9]: La, S., Yoo, H. H., & Kim, D. (2005). Liquid Chromatography−Mass Spectrometric Analysis of Urinary Metabolites and Their Pattern Recognition for the Prediction of Drug-Induced Hepatotoxicity. *Chemical Research in Toxicology, 18*(12), 1887-1896

[10]: Yamazaki, H. (2016). Differences in Toxicological and Pharmacological Responses Mediated by Polymorphic Cytochromes P450 and Related Drug-Metabolizing Enzymes. *Chemical Research in Toxicology, 30*(1), 53-60.